Г.Е. ШИЛОВ

ОСНОВЫ СОВРЕМЕННОГО АНАЛИЗА

(Методическое пособие)

Выпуск I.

ПРОИЗВОДНАЯ

МГУ - 1974

ВВЕДЕНИЕ

Математическим анализом называют обычно большую область математики, связанную с понятиями функции, производной, интеграла. В нее входит ряд меньших областей - дифференциальное и интегральное исчисление, дифференциальные уравнения - обыкновенные и с частными производными, дифференциальная геометрия, интегральные уравнения и другие. Существуют еще термины"классический математический анализ" и "современный математический анализ". Было бы трудно дать этим названиям четкие определения. Но приблизительный смысл этих слов следующий. В классическом анализе операции дифференцирования и интегрирования строятся применительно к функциям одного или нескольких вещественных переменных с числовыми значениями; в современном анализе эти операции устанавливаются для объектов значительно более общей природы - например, функций, аргументы и значения которых лежат в нормированных пространствах. Такое обобщение появилось не случайно и не по прихоти каких-то ученых. Оно дало возможность установить новые глубокие связи между различными областями анализа; классический анализ с точки зрения современного анализа приобретает значительно большую стройность и единство, чем ранее, и способность к новым широким применениям. С новым общим понятием производной тесно связан и новый объект анализа- гладкие многообразия, абстрактное определение которых на основе классических работ Римана, Пуанкаре, Леви-Чивита выкристаллизовалось в 30-х гг. нашего века (Уитни). Локальное устройство гладкого многообразия, с точки зрения дифференциальных операций, таково же, как у линейного n -мерного или нормированного - пространства. Но, обогащая гладкое многообразие дополнительными структурами - римановой метрикой, симплектической геометрией, связностью, групповой операцией и др. - даже в локальных вопросах мы чувствовали бы стеснение, оставаясь в рамках линейного пространства; а рассмотрение глобальных задач здесь открывает пути для новых глубоких и содержательных исследований, связанных разнообразными нитями со многими математическими и смежными дисциплинами. Так, понятия и методы современного анализа широко используются в аналитической механике, теории вероятностей, математической экономике и др.

Предлагаемое методическое пособие выходит выпусками, в каждом по одной главе. Для удобства ссылок имеется единая рубрикация. Необходимые ссылки на элементарную теорию по большей части даются по книге автора "Математический анализ, Функции одного переменного" с первой цифрой 0; 012.45а означает эту книгу, гл.12, § 4, пункт 45а. В этих выпусках имеются некоторые пересечения с книгой автора "Функции нескольких вещественных переменных, чч.1-2", но даже в таких местах здесь дается в значительной мере новое освещение вороса и новые применения.

Глава I

ПРОИЗВОДНАЯ

§ I.I. Функции.

I.IIa. Пусть дана функция $y = f(x)$, определенная на некотором множестве X и принимающая свои значения в множестве Y. В дальнейшем, в зависимости от целесообразности, будем употреблять следующие формы записи этого факта:

$$y : X \to Y \; ; \quad y = f(x) : X \to Y \; ; \; y = f(x) \; (X \to Y) \; ;$$

$$x \to y = f(x) \; ; \quad x \to f(x) \; ;$$

последние две - в случаях, когда множества X и Y известны из контекста. Если необходимо указать, что функция $f(x)$ определена на подмножестве $E \subset X$ и принимает значения в подмножестве $F \subset Y$ будем также писать

$$y = f(x) : (E \subset X) \to (F \subset Y),$$

или, опуская скобки, $y=f(x): E\subset X \to F\subset Y.$

Две функции $y=f(x): X\to Y$ и $g(x): X\to Y$ считаются тождественно равными – в обозначении $f(x)\equiv g(x)$, – или $f=g$ – тогда и только тогда, когда для каждого $x\in X$ элементы $f(x)\in Y$ и $g(x)\in Y$ совпадают. Если хотя бы для одного $x=x_0\in X$ это равенство не имеет места, т.е. $f(x_0)\neq g(x_0)$, функции $f(x)$ и $g(x)$ считаются различными.

Всегда определена тождественная функция $e_X: X\to X$, действующая по правилу $e_X x = x$ для любого $x\in X$.

б. Функция $f(x)$ называется числовой, если $Y\subset R_1$ (вещественная ось). Такую функцию $f(x)$ называют также вещественной (точнее, вещественно-значной). Если $Y\not\subset R_1$, функция $f(x)$ называется, как правило, отображением. Если Y есть линейное пространство , например, если $Y=R_n$ (n -мерное вещественное пространство), то функция $f(x): X\to Y$ называется векторной. Если X – множество в R_1, то $f(x)$ называется функцией одного вещественного переменного. Если X есть область в R_n, то $f(x)$ называется функцией n вещественных переменных; этими переменными считаются обычно координаты $x_1,\dots,x_n$ точки x в каком-либо базисе пространства R_n.

в. Последнее определение можно следующим образом обобщить. Пусть $X_1,\dots,X_n$ некоторые множества; тогда совокупность X всех наборов

$$x=\{x_1,\dots,x_n\},\ x_1\in X_1,\dots,\ x_n\in X_n \qquad (I)$$

называется прямым произведением множеств $X_1,\dots,X_n$ и обозначается так:

$$X=X_1\times\dots\times X_n=\prod_{i=1}^{n} X_i \qquad (2)$$

В выражении (I) элементы $x_1,\dots,x_n$ называются составляющими (или координатами) элемента x. Пусть далее имеется функция $y=f(x): E\subset X\to Y$. Поскольку каждый элемент $x\in E$

определяется заданием n своих координат $x_1, \ldots, x_n$, функция $f(x)$ может быть рассматриваема, как функция от n переменных $x_1, \ldots, x_n$, пробегающих множество E. Разумеется, в общем случае эти переменные уже не являются вещественными.

I.12. Графики, годографы, поверхности уровня.

а. Чтобы наглядно представить себе числовую функцию одного вещественного переменного, мы рисовали ее график, откладывая в каждой точке x ее области определения на вещественной оси значение функции $f(x)$ по направлению оси y.

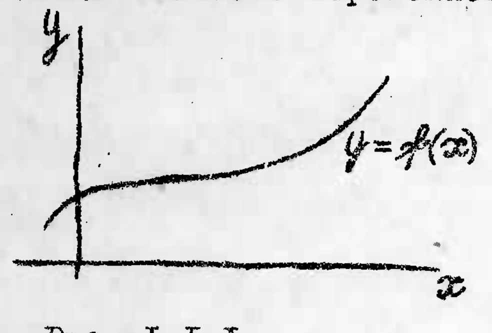


Рис. I.I-I

В случае числовой функции двух вещественных переменных $f(x_1, x_2)$ можно в каждой точке (x_1, x_2) множества X на плоскости R_2 откладывать значение $f(x_1, x_2)$ в направлении третьей оси y. В случае числовой функции одного переменного ее графиком служит, вообще говоря, некоторая кривая; в случае двух переменных график числовой функции будет представлять собою, по крайней мере для простых функций, некоторую поверхность, которую можно изобразить на чертеже, пользуясь правилами перспективы.

П р и м е р ы. Линейной функции $y = k_1 x_1 + k_2 x_2 + b$ в качестве графика отвечает некоторая плоскость. Числа

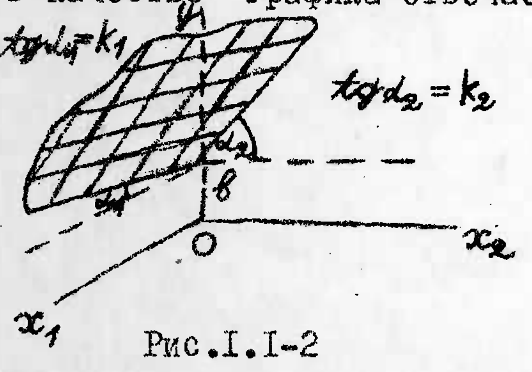


Рис.I.I-2

Рис. I.I-3

k_1 и k_2 называются угловыми коэффициентами этой плоскости, а их геометрический смысл очевиден из рис. I.I-2.

Графиком квадратичной функции $y = x_1^2 + x_2^2$ является параболоид вращения (при $y = \alpha_1 x_1^2 + \alpha_2 x_2^2$, $\alpha_1 > 0, \alpha_2 > 0$ – эллиптический параболоид), показанный на рис. I.I-3.

Графиком квадратичной функции $y = x_2^2 - x_1^2$ является седлообразная поверхность, называемая гиперболическим параболоидом; если ось y направлена вверх, то вертикальные сечения поверхности – параболы, а горизонтальные сечения – гиперболы.

/рисунок 1.1-4/

Рис. 1.1-4

В общем случае графиком функции $y=f(x): E\subset X \to Y$ по определению, является множество точек $\{x, f(x)\}$ в прямом произведении X и Y. Так, в случае числовой функции $f(x_1,\dots,x_n)$ от n переменных /трех и более/ ее графиком является совокупность точек $\{x_1,\dots,x_n, f(x_1,\dots,x_n)\}$ в $n+1$-мерном пространстве.

Однако наглядное представление этого "графика" затруднительно; в таких случаях наглядность должна уступать место логике. Впрочем, графики не необходимы и в случае двух переменных или даже одного переменного, хотя никто не отрицает их полезности.

б. Более простой, чем график функции $y=f(x): E\subset X\to Y$, хотя и менее содержательной геометрической иллюстрацией является указание годографа, т.е. множества $f(E)$ всех значений, принимаемых функцией $f(x)$ на ее области определения E. Для числовой функции $y=f(x): E\subset X\to R_1$ множество $f(E)$ есть некоторое подмножество вещественной оси, и, вообще говоря, мало что может дать для изучения самой функции $f(x)$. Более содержательная информация о функции $f(x)$ по множеству $f(E)$ получается, когда размерность пространства Y больше, чем размерность пространства X. Рассмотрим, например, функцию одного вещественного переменного $y=f(x): E\subset R_1\to Y$; здесь множество $f(E)$, вообще говоря, есть некоторая кривая в пространстве Y, и ее вид уже содержит некоторую информацию. Так, для функции $y=(y_1,y_2) = f(\varphi): E=\{0\le\varphi<2\pi\}\subset R_1\to R_2$ заданной уравнениями

$$y_1 = r\cdot\cos\varphi$$
$$y_2 = r\cdot\sin\varphi$$

множество $f(E)$ представляет собою окружность радиуса r с уравнением $y_1^2+y_2^2=r^2$. Параметр φ допускает истолкование, как полярный угол (рис. I.I-5). Для достаточно простой функции двух вещественных переменных $y=f(x): E\subset R_2\to Y$ множество $f(E)$ представляет собою двумерную поверхность в пространстве Y. Так, для функции $y=(y_1,y_2,y_3)=f(\theta,\varphi)$:

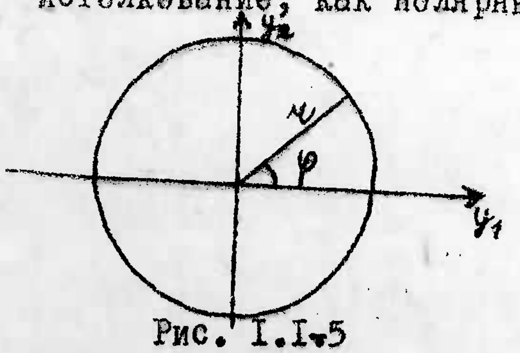


Рис. I.I-5

$E=\{0\le\theta\le\pi\}\times\{0\le\varphi\le 2\pi\}\subset R_2\to R_3$,

заданной уравнениями $y_1=r\cdot\sin\theta\cos\varphi$,

$y_2=r\cdot\sin\theta\cdot\sin\varphi$, $y_3=r\cdot\cos\theta$.

множество $f(E)$ представляет собою сферу

$$y_1^2+y_2^2+y_3^2=r^2$$

И здесь параметры θ, φ допускают геометрическое истолкование, как "сферические углы" (рис. I.I-6). В общем случае координаты x в пространстве X не имеют геометрического истолкования в $f(E)$.

Тем не менее, следует иметь в виду, что годограф, как геометрическое место точек в Y, не определяет однозначно саму функцию $f(x)$; могут быть различные функции $G\to Y$ с одним и тем же годографом. Чтобы по годографу можно было восстановить функцию, нужно для каждой его точки y указать, из какой точки $x\in G$ (или из каких точек $x\in G$ - для данной точки y таких x может быть несколько) получается эта точка y применением функции f.

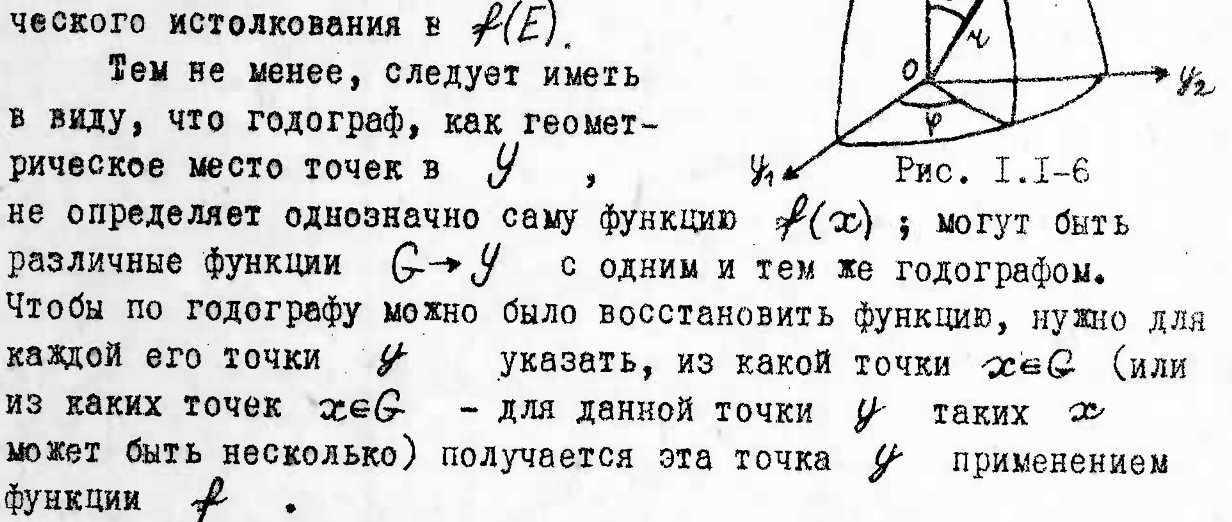


Рис. I.I-6

в. В некоторых случаях наглядное представление о функции можно получить из рассмотрения ее линий (или поверхностей) уровня. Линия (поверхность) уровня функции $y=f(x)$ есть геометрическое место точек, где функция сохраняет какое-либо постоянное значение $y=y_0$. Так, для функции

$f(x)= \rho(x,a)=|x-a|$ $(R_2\to R_1)$ линии уровня (рис. I.I-7) суть окружности с центром в точке a (а также и сама точка a , где функция $f(x)$ принимает значение 0). Для функции $f(x)=\rho(x,a)+\rho(x,b)$ $(R_2\to R_1)$ линии уровня представляют собою (рис.I.I-8) эллипсы с фокусами в точках a и b (и отрезок, соединяющий точки a и b); для функции $f(x)=\rho(x,a)-\rho(x,b)$ (рис. I.I-9) – гиперболы с фокусами в a и b (включая прямую – ось симметрии и две полупрямых); для функции $f(x)=\rho(x,a)\cdot\rho(x,b)$ $(R_2\to R_1)$ – семейство (рис.I.I-10) овалов Кассини (среди них лемниската Бернулли, см. задачу 6); для функции $f(x)=\rho(x,a)/\rho(x,b)$ $(R_2\to R_1)$ – некоторое семейство окружностей с центрами на прямой ab и ось симметрии точек a и b (рис. I.I-II). Поверхности уровня тех же функций, рассматриваемых в R_3 , возникают от вращения получающихся кривых вокруг оси ab .

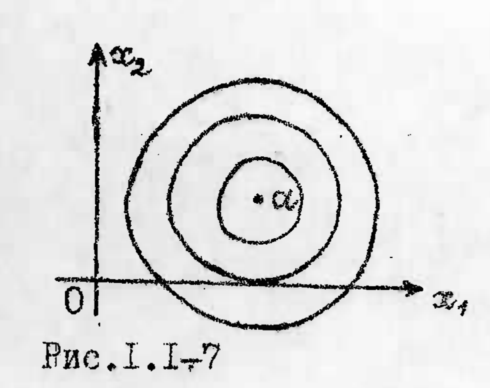
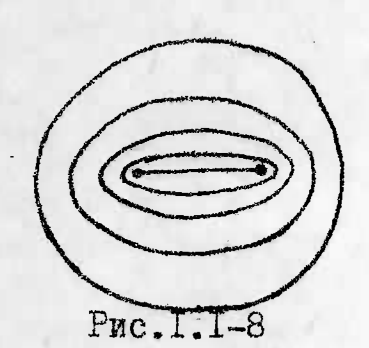


Рис.I.I-7


Рис.I.I-8

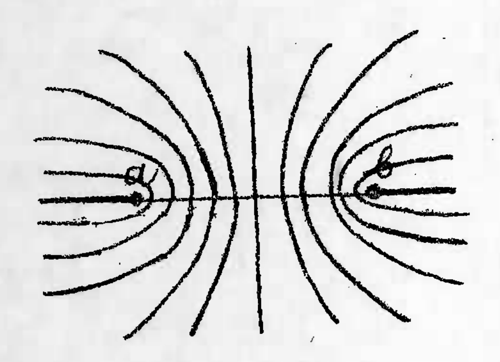


Рис.I.I-9

Рис.I.I-10

Естественно, что существуют тесные связи между графиком функции $y=f(x): E\subset X\to Y$ (расположенным в пространстве $X\times Y$), ее годографом (в пространстве Y) и поверхностями уровня (в пространстве X). Заметим, кроме того, что график функции $y=f(x): E\subset X\to Y$ есть в то же

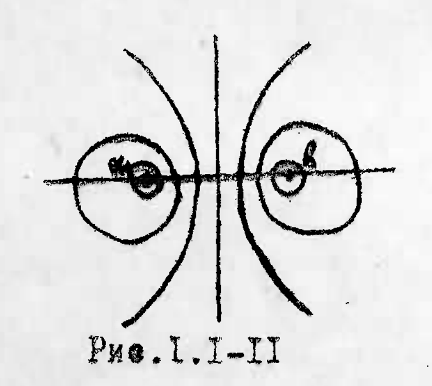


Рис.I.I-II

время множество значений функции $\mathcal{F}(x)=\{x, f(x)\}: E\subset X\to X\times Y$, а также одна из поверхностей уровня функции $\Phi(x,y)=y-f(x): Y\times E\to Y$ (именно, $\Phi(x,y)=0$).

I.I3. Композиция функций.

а. Пусть имеются три множества X, Y, Z и даны функции $y=\varphi(x): X\to Y$ и $z=f(y): Y\to Z$. Можно образовать новую функцию $z=f[\varphi(x)]: X\to Z$; она называется сложной функцией (составленной из φ и f) или композицией функций f и φ; иногда композиция функций f и φ обозначается символом $f\circ\varphi$. (Заметим, что в описанной обстановке символ $\varphi\circ f$ не имеет смысла, так как значения функции f лежат в Z, а функция φ определена на X). Пусть $X=Z$, так что $\varphi\circ f$ и $f\circ\varphi$ определены; природа их остается различной, так как $\varphi\circ f$ действует из Y в X, а $f\circ\varphi$ - из X в Y. Если при этом $X=Y$, то обе композиции $f\circ\varphi$ и $\varphi\circ f$ действуют из X в X; тем не менее, они могут быть различными (см. задачу 32).

Для отображения $y=f(x): X\to Y$ отметим, кроме того, очевидные равенства
$$f\circ e_X=f,\ e_Y\circ f=f,$$
где e_X и e_Y соответствующие тождественные отображения (I.IIа).

б. Ассоциативность композиции. Пусть имеются четыре множества X, Y, Z, W и три функции: $y=\varphi(x): X\to Y$, $z=f(y): Y\to Z$, $w=g(z): Z\to W$.

Композиция $f\circ\varphi$ действует из X в Z, поэтому определена композиция $g\circ(f\circ\varphi)$, действующая из X в W. С другой стороны, композиция $g\circ f$ действует из Y в W,

поэтому определена композиция $(g\circ f)\circ \varphi$, также действующая из X в W. Утверждается, что имеет место равенство

$$g\circ(f\circ\varphi)=(g\circ f)\circ\varphi$$

В соответствии с определением равенства функций (I.IIa) достаточно проверить, что для любого $x\in X$

$$[g\circ(f\circ\varphi)](x)=[(g\circ f)\circ\varphi](x) \qquad (I)$$

Левая часть раскрывается, согласно определению композиции, так:

$$[g\circ(f\circ\varphi)](x)=g[(f\circ\varphi)(x)]=g[f(\varphi(x))] \qquad (2)$$

правая часть, в силу того же определения, так:

$$[(g\circ f)\circ\varphi](x)=(g\circ f)\,\varphi(x)=g[f(\varphi(x))] \qquad (3)$$

Правые части в (3) и (2) совпадают, что нам и требуется. Равенство (I) выражает закон ассоциативности композиции функций.

в. Обратная функция. Пусть X и Y два множества и имеются три функции $y=f(x): X\to Y$, $x=\varphi(y): Y\to X$, $x=\psi(y): Y\to X$. Пусть далее, как обычно, e_X есть тождественное отображение $X\to X$ и e_Y - тождественное отображение $Y\to Y$.

Композиции $f\circ\varphi$ и $f\circ\psi$ действуют из Y в Y, композиции $\varphi\circ f$ и $\psi\circ f$ из X в X.

Если выполняется равенство

$$\varphi\circ f=e_X \qquad (4)$$

отображение φ называется левым обратным для f, а f - правым обратным для φ; соответственно, если

$$f\circ\psi=e_Y \qquad (5)$$

то ψ называется правым обратным для f, а f - левым обратным для ψ. Для данного f может не существовать ни правого, ни левого обратных, может существовать много правых (или левых обратных) (зад.33); оказывается, однако, что если у отображения f имеется левое обратное φ и имеется правое обратное ψ, то $\varphi=\psi$ и других левых или правых обратных нет.

Действительно, пусть выполнены равенства (4) и (5). Произведем в равенстве (4) композицию справа с ψ: мы получим

$$(\varphi \circ f) \circ \psi = e_X \cdot \psi = \psi$$

С другой стороны, в силу ассоциативности композиции и на основании (4)

$$(\varphi \circ f) \circ \psi = \varphi \circ (f \circ \psi) = \varphi \circ e_y = \varphi$$

так что $\varphi = \psi$. Если φ_1 другое левое обратное к f, то мы имеем $\varphi_1 = \psi = \varphi$, так что $\varphi_1 = \varphi$; аналогично любое правое обратное ψ_1 к f совпадает с ψ, что и утверждалось.

Отображение f, имеющее левое и правое обратное, называется обратимым, а его левое - и правое - обратное, по доказанному, единственное, называется обратным к f. Отображение $f(x): X \to y$ тогда и только тогда обратимо, когда оно является взаимно-однозначным с X на все y.

Действительно, если отображение $y = f(x)$ обратимо и $x = \varphi(y)$ обратное отображение, и если $f(x_1) = f(x_2) = y$, то $\varphi(y) = \varphi(f(x_1)) = x_1$ и в то же время $\varphi(y) = \varphi(f(x_2)) = x_2$, так что $x_1 = x_2$, поэтому функция f отображает X взаимно однозначно (может быть, еще не на все y). С другой стороны, для заданного $y_0 \in y$ положим $\varphi(y_0) = x_0$; отсюда $f(x_0) = f(\varphi(y_0)) = y_0$, так что функция f отображает X на все y.

С другой стороны, если отображение $y = f(x): X \to y$ взаимно однозначно (с X на все y), то естественно определено отображение $\varphi(y): y \to X$, ставящее в соответствие заданному $y \in y$ то единственное $x \in X$, для которого $f(x) = y$; таким образом, $\varphi(f(x)) = x$ и $f(\varphi(y)) = y$, так что отображение φ является обратным к f.

I.I4. Линейные функции.

а. Чтобы иметь возможность говорить о линейной функции $A(x): X \to y$, мы должны предположить, что X и y - линейные пространства для определенности, над полем вещественных чисел.

Напомним, что множество X называется линейным пространством над полем вещественных чисел, если в X определены

операции сложения любых двух элементов и умножения элементов на вещественные числа с выполнением обычных свойств этих операций . В частности, в пространстве X имеется нуль - такой элемент 0, что $x+0=x$ для любого $x \in X$.

Наличие линейных операций позволяет ввести в X понятия аффинной геометрии. Прямая, проходящая через точки a и b есть совокупность всех точек вида $(1-t)a+tb$, где $t \in R_1$; если при этом t меняется лишь на $[0,1]$, мы получаем в X отрезок, соединяющий точки a и b. Значению $t=0$ отвечает, очевидно, точка a, значению $t=1$ - точка b. Можно записать уравнение этой же прямой в виде $a+tc$, где $c=b-a$; в такой записи a называют начальной точкой, а вектор $c=b-a$ - направляющим вектором прямой. Множество $L \subset X$, содержащее вместе с каждыми двумя точками a и b всю проходящую через них прямую, называется линейным многообразием в X. Множество $Q \subset X$, содержащее вместе с каждыми двумя точками a и b соединяющий их отрезок, называется выпуклым в X. Линейное многообразие L, содержащее 0, является подпространством в X; если же линейное многообразие L не содержит 0, оно получается из некоторого подпространства сдвигом (т.е. прибавлением ко всем точкам фиксированного вектора).

6. Пусть X и Y линейные пространства над полем вещественных чисел.

Функция $A(x): X \to Y$ называется линейной функцией, или линейным оператором, если для любых x_1 и x_2 из X и любых вещественных α_1 и α_2 выполнится равенство

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A(x_1) + \alpha_2 A(x_2)$$

Из этого определения следует, что линейный оператор $A(x)$ сохраняет аффинные свойства геометрических объектов: прямую переводит в прямую (или в точку), отрезок - в отрезок, линейное многообразие - в линейное многообразие, выпуклое множество - в выпуклое множество.

Для примера рассмотрим n -мерное линейное пространство $X = R_n$ и фиксируем в нем базис. Сопоставляя каждой точке $x \in R_n$ какую-либо ее координату x_k $(k=1,\dots,n)$, мы получаем линейную числовую функцию в пространстве R_n. Наиболее общей числовой линейной функцией в R_n является функция $A(x) = \sum_{k=1}^{n} c_k x_k$, где c_k $(k=1,\dots,n)$ – заданные постоянные числа.

Пусть далее $Y = R_m$ m – мерное пространство, так что в некотором фиксированном базисе каждый вектор $y \in Y$ получает координаты $y_1,\dots,y_m$. Пусть $\|a_{ik}\|$ – числовая матрица с m строками и n столбцами (короче, $m\times n$ – матрица). Тогда система равенств

$$y_i = \sum_{k=1}^{n} a_{ik} x_k ,\ i=1,\dots,m \qquad (I)$$

определяет линейную функцию $y = A(x) : X \to Y$. Известно, что любая линейная функция из R_n в R_m может быть записана в виде (I).

в. Тождественный оператор e_X, действующий по правилу $e_X x \equiv x$, является, очевидно, линейным; он обозначается через E_X или просто через E, если X известно.

Если линейный оператор $A: X \to Y$ обратим, то обратное отображение $B: Y \to X$ есть также линейный оператор. Действительно, обратимое отображение A, как мы видели в I.I3 в, осуществляет взаимно однозначное отображение X на Y; поэтому для любых $y_1 \in Y$ и $y_2 \in Y$ можно найти также $x_1 \in X$, $x_2 \in X$, что $y_1 = A x_1$, $y_2 = A x_2$; тогда $B y_1 = x_1$, $B y_2 = x_2$ и

$$B(\alpha_1 y_1 + \alpha_2 y_2) = B(\alpha_1 A x_1 + \alpha_2 A x_2) =$$

$$= BA(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 x_1 + \alpha_2 x_2 = \alpha_1 B y_1 + \alpha_2 B y_2 ,$$

что и требуется.

Этот обратный оператор обозначается через A^{-1}.

г. Для линейных операторов $A: X \to Y$ и $B: X \to Y$ и любых чисел α и β определена линейная комбинация $\alpha A + \beta B$, как линейный непрерывный оператор, переводящий X в Y по формуле $(\alpha A + \beta B)x = \alpha A(x) + \beta B(x)$.

Таким образом, совокупность всех линейных операторов, действующих из X в Y, сама представляет собою линейное пространство: оно обозначается через $L(X,Y)$.

Если $X=R_1$, то $L(X,Y)=L(R_1,Y)$ естественно отождествляется с самим Y: любому $a\in Y$ соответствует оператор $A\in L(R_1,Y)$, действующий по формуле $At=t\cdot a\ (t\in R_1)$, и любой оператор $A\in L(R_1,Y)$ действует по формуле $At=At\cdot 1=t\cdot A(1)=ta$, где $a=A(1)$. Пространство $L(X,X)$ обозначается короче через $L(X)$.

Если $Y=R_1$, то соответствующий линейный оператор $A: X\to R_1$, как правило, называется линейным функционалом. Если X имеет размерность n, то и $L(X,R_1)$ имеет размерность n.

д. Для линейных операторов $A: X\to Y$ и $B: Y\to Z$ их композиция $C=BA: X\to Z$, определенная по общему правилу $Cx=B(Ax)$, есть снова линейный оператор.

е. Прямая сумма линейных пространств. Пусть имеются линейные пространства $X_1,\ldots,X_n$ и $X=X_1\times\ldots\times X_n$ есть прямое произведение множеств $X_1,\ldots,X_n$ (I.IIв), т.е. совокупность комплексов $x=\{x_1,\ldots,x_n\}$, $x_1\in X_1,\ldots,x_n\in X_n$. Введем в X линейные операции по правилам

$$\{x_1,\ldots,x_n\}+\{y_1,\ldots,y_n\}=\{x_1+y_1,\ldots,x_n+y_n\}$$

$$\alpha\{x_1,\ldots,x_n\}=\{\alpha x_1,\ldots,\alpha x_n\}$$

Нетрудно проверить, что эти операции превращают X в линейное пространство; в этом случае оно называется прямой суммой пространств $X_1,\ldots,X_n$. Составляющие $x_1,\ldots,x_n$ есть функции от x, которые мы обозначим так:

$$x_i=P_i(x),\quad i=1,\ldots,n$$

Функция $P_i(x)$ является линейным оператором из X в X_i; он называется проектором (или оператором проектирования) из X в X_i. Сами пространства X_i можно естественно отождествить с некоторыми подпространствами $X_i'\subset X$;

именно, в качестве X_i' можно взять совокупность элементов вида

$$\{0,\ldots,0,x_i,0,\ldots,0\},\ x_i\in X_i$$

Подпространства X_i' обладают следующими свойствами:
(α) для каждого $x\in X$ имеется разложение

$$x=x_1'+\cdots+x_n',\ x_k'\in X_k' \qquad (2)$$

(β) это разложение единственно, т.е. из (2) и

$$x=x_1''+\cdots+x_n'',\ x_k''\in X_k''$$

следует, что $x_1'=x_1'',\ldots,x_n'=x_n''$.

Действительно, мы имеем $x=\{x_1,\ldots,x_n\}=$ $=\{x_1,0,\ldots,0\}+\cdots+\{0,\ldots,0,x_n\}$, что дает требуемое разложение (2); если же, наряду с этим, имеется разложение

$$x=\{\tilde{x}_1,0,\ldots,0\}+\cdots+\{0,\ldots,0,\tilde{x}_n\}=\{\tilde{x}_1,\ldots,\tilde{x}_n\},$$

то по определению X мы имеем $x_1=\tilde{x}_1,\ldots,x_n=\tilde{x}_n$.

Обратно, если в пространстве X выделены каким-то образом подпространства $X_1',\ldots,X_n'$ так, что выполняются условия $(\alpha)-(\beta)$, то прямая сумма X пространств $X_1',\ldots,X_n'$ изоморфна пространству X. Именно данному $x'=\{x_1',\ldots,x_n'\}$ поставим в соответствие элемент $x=x_1'+\cdots+x_n'\in X$; из условий $(\alpha)-(\beta)$ легко следует, что соответствие $x'\to x$ является изоморфизмом. В этом случае будем говорить, что пространство X разложено в прямую сумму подпространств $X_1',\ldots,X_n'$.

ж. Пусть пространство X есть прямая сумма подпространств $X_1,\ldots,X_n$, так что для любого $x\in X$ имеем $x=\mathcal{P}_1x+\cdots+\mathcal{P}_nx$, где $\mathcal{P}_1,\ldots,\mathcal{P}_n$ – соответствующие проекторы; пусть, аналогично, пространство $\mathcal{Y}$ есть прямая сумма подпространств $\mathcal{Y}_1,\ldots,\mathcal{Y}_m$, так что для любого $y\in\mathcal{Y}$ имеем $y=Q_1y+\cdots+Q_my$, где $Q_1,\ldots,Q_m$ соответствующие проекторы. Пусть $A:X\to\mathcal{Y}$ линейный оператор. Положим

$$A_{ij}=Q_jA\mathcal{P}_i\ (i=1,\ldots,n,\ j=1,\ldots,m) \qquad (3)$$

Операторы A_{ij} действуют из X в $\mathcal{Y}_i$, но естественно рассматривать оператор A_{ij} заданным только на X_j. Эти операторы составляют операторную $n\times m$-матрицу

$$\|A_{ij}\| \equiv \begin{Vmatrix} A_{11} & \cdot & \cdot & A_{1n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ A_{m1} & \cdot & \cdot & A_{mn} \end{Vmatrix} \qquad (4)$$

Для любого $x = \sum_{i=1}^{n} P_i x \equiv \sum_{i=1}^{n} x_i \equiv \sum_{i=1}^{n} P_i x_i$ имеем

$$y_j = Q_j y = Q_j A x = \sum_{i=1}^{n} Q_j A P_i x = \\ = \sum_{i=1}^{n} Q_j A P_i x_i = \sum_{i=1}^{n} A_{ij} x_i , \qquad (5)$$

что напоминает обычную матричную запись линейного оператора, действующего из R_n в R_m. Формула (5) показывает, что оператор A полностью определен матрицей $\|A_{ij}\|$.
Более того, любому наперед заданному набору линейных операторов $A_{ij} : X_i \to Y_j \; (i=1,..,n;\ j=1,...,m)$ отвечает линейный оператор $A : X \to Y$, действующий по формуле

$$A x = \sum_{j=1}^{m} \sum_{i=1}^{n} A_{ij} x_i ,$$

или, что то же,

$$y_j = (A x)_j = \sum_{j=1}^{n} A_{ij} x_i ;$$

легко проверить, что соответствующая ему операторная $n \times m$ -матрица, построенная по указанному выше правилу, совпадает с матрицей $\|A_{ij}\|$. Таким образом, операторы $A : X \to Y$ находятся в естественном взаимно-однозначном соответствии с операторными $n \times m$ -матрицами $\|A_{ij}\|$.
В частности, для $Y = X$, тождественному линейному оператору $E : X \to X$, $E x = x$, очевидно, отвечает матрица

$$\begin{Vmatrix} E_1 & & \\ & \ddots & \\ & & E_n \end{Vmatrix}$$

(невыписанные элементы матрицы - нулевые операторы), где E_i есть тождественный оператор в подпространстве $X_i \; (i=1,..., n)$.

Далее заметим, что для двух линейных операторов

$A: X \to Y$ и $B: X \to Y$, с соответствующими матрицами $\|A_{ij}\|$ и $\|B_{ij}\|$, линейной комбинации $\alpha A + \beta B$ отвечает, очевидно, матрица $\|\alpha A_{ij} + \beta B_{ij}\|$. Поэтому указанное нами взаимно-однозначное соответствие является, сверх того, изоморфизмом линейного пространства $L(X,Y)$ и линейного пространства всех операторных $n \times m$ - матриц $\|A_{ij}\|$ (с естественными линейными операциями).

В частности, при предположениях е пространство $L(X,Y)$ оказывается изоморфным прямой сумме $m \cdot n$ пространств $L(X_i, Y_j)$. Беря $n=1$, получаем, что $L(X,Y)$ изоморфно прямой сумме m пространств $L(X,Y_j)$, а беря $m=1$ - что $L(X,Y)$ изоморфно прямой сумме n пространств $L(X_i,Y)$.

з. Пусть, кроме пространств X и Y из е, имеется третье пространство Z, являющееся прямой суммой подпространств $Z_1 + \dots + Z_p$. Пусть дан линейный оператор $B: Y \to Z$; по имеющимся разложениям $Y = Y_1 + \dots + Y_m$ и $Z = Z_1 + \dots + Z_p$ ему можно поставить в соответствие матрицу $\|B_{jk}\|$, $j = 1, \dots, m$; $k = 1, \dots, p$.

Композиция $C = BA$ есть линейный оператор, действующий из X в Z; при этом с одной стороны,

$$z_i = \sum_k C_{ik} x_k \,, \quad C_{ik}: X_k \to Z_i$$

, а с другой стороны

$$z_i = \sum_j B_{ij} y_j = \sum_{j,k} B_{ij} A_{jk} x_k .$$

Полагая здесь $x_s = 0$ при $s \neq k$, а $x_k \in X_k$ - произвольными, получаем

$$C_{ik} = \sum_{j=1}^{m} B_{ij} A_{jk}$$

Это правило соответствует обычному правилу умножения матриц.

и. Пусть $Z = X$, $Z_i = X_i$ и оператор B является левым обратным для A, так что $BA = E_X$. Это равенство соответствует системе

$$\sum_{j=1}^{m} B_{ij} A_{jk} = \begin{cases} 0 & \text{при } i \neq k, \\ E_k & \text{при } i = k \ (=1,\dots,n), \end{cases}$$

где E_k - тождественный оператор в пространстве X_k.
Если при этом B является и правым обратным для A,
то $A \cdot B = E_y$ или

$$\sum_{j=1}^{m} A_{ij} B_{jk} = \begin{cases} 0 \text{ при } i \neq k \\ E'_k \text{ при } i = k \; (=1, \ldots, m) \end{cases}$$

где E'_k - тождественный оператор в пространстве y_k.
к. Как по матрице $\|A_{ij}\|$ узнать, является ли оператор A обратимым ? В конечномерном случае ответ дается необходимым и достаточным условием $\det \|A_{ij}\| \neq 0$.
В общем случае имеются лишь достаточные условия; рассмотрим для простоты случай $m = n = 2$, $X = y$.

В дальнейшем операторы $A_{ij}, B_{ij}, C_{ij}, S_{ij}, T_{ij}$ действуют из подпространства X_j в X_i $(i, j = 1, 2)$

ЛЕММА. Обратимость операторов A_{11} и $D_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12}$ является достаточным условием для разрешимости системы

$$A_{11} B_{11} + A_{12} B_{21} = S_{11} \quad (1)$$

$$A_{21} B_{11} + A_{22} B_{21} = S_{21} \quad (2)$$

и для разрешимости системы

$$C_{11} A_{11} + C_{12} A_{21} = T_{11} \quad (3)$$

$$C_{11} A_{12} + C_{12} A_{22} = T_{12} \quad (4)$$

(A_{ij}, S_{ij}, T_{ij} - заданы, B_{ij}, C_{ij} - искомые).

Доказательство. Допустим на момент, что решение системы (1)-(2) существует и дадим его явное выражение. Для этого умножим (1) слева на A_{11}^{-1}; получим

$$B_{11} + A_{11}^{-1} A_{12} \cdot B_{21} = A_{11}^{-1} \cdot S_{11} \quad (5)$$

Это равенство умножим слева на A_{21} и затем используем (2); мы получим

$$-A_{22} B_{21} + A_{21} A_{11}^{-1} A_{12} B_{21} = A_{21} A_{11}^{-1} \cdot S_{11} - S_{21}$$

или $(-A_{22} + A_{21} A_{11}^{-1} A_{12}) \cdot B_{21} = A_{21} \cdot A_{11}^{-1} \cdot S_{11} - S_{21}$

Умножая на $\mathcal{D}_{22}^{-1}$, получаем

$$B_{21} = -\mathcal{D}_{22}^{-1} \cdot A_{21} A_{11}^{-1} S_{11} + \mathcal{D}_{22}^{-1} \cdot S_{21} \tag{6}$$

откуда и из (5)

$$B_{11} = A_{11}^{-1} S_{11} + A_{11}^{-1} A_{12} \mathcal{D}_{22}^{-1} A_{21} A_{11}^{-1} S_{11} - A_{11}^{-1} A_{12} \mathcal{D}_{22}^{-1} S_{21} \tag{7}$$

Теперь, не предполагая разрешимости системы (I)-(2), подставим в нее B_{11} и B_{21} из формул (6) и (7). Мы увидим, что эта система удовлетворится; таким образом, для системы (I)-(2) утверждение леммы доказано. Для системы (3)-(4) действуем аналогично: предполагая существование решения, приходим к уравнениям

$$C_{11} + C_{12} A_{21} \cdot A_{11}^{-1} = T_{11} \cdot A_{11}^{-1}$$

$$-C_{12} A_{22} + C_{12} A_{21} \cdot A_{11}^{-1} \cdot A_{12} = T_{11} \cdot A_{11}^{-1} \cdot A_{12} - T_{12}$$

$$C_{12} = -T_{11} \cdot A_{11}^{-1} \cdot A_{12} \cdot \mathcal{D}_{22}^{-1} + T_{12} \cdot \mathcal{D}_{22}^{-1} \tag{8}$$

$$C_{11} = T_{11} A_{11}^{-1} + T_{11} A_{11}^{-1} A_{12} \mathcal{D}_{22}^{-1} A_{21} A_{11}^{-1} - T_{11} \cdot A_{11}^{-1} \tag{9}$$

Прямой проверкой убеждаемся, что (8) и (9) дают решение системы (3)-(4), и лемма доказана.

Теорема. Обратимость операторов A_{11} и $\mathcal{D}_{22} = A_{22} - A_{21} \cdot A_{11}^{-1} \cdot A_{12}$ является достаточным условием для обратимости оператора A, соответствующего матрице

$$\begin{Vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Vmatrix}$$

Доказательство. Обратимость оператора A справа эквивалентна разрешимости системы уравнений

$$A_{11} B_{11} + A_{12} B_{21} = E_{11},$$

$$A_{21} B_{11} + A_{22} B_{21} = O,$$

$$A_{11}B_{12}+A_{12}B_{22}=O,$$
$$A_{21}B_{12}+A_{22}B_{22}=E_{22}.$$

Эта система разбивается на две независимые системы типа (I)-(2): в первой системе искомыми являются B_{11} и B_{21}, во второй — B_{12} и B_{22}. Условие теоремы, согласно лемме, обеспечивает разрешимость обеих этих систем. Аналогично, обратимость оператора A слева эквивалентна разрешимости системы уравнений

$$C_{11}A_{11}+C_{12}A_{21}=E_{11}$$
$$C_{11}A_{12}+C_{12}A_{22}=O,$$
$$C_{21}A_{11}+C_{22}A_{21}=O,$$
$$C_{21}A_{12}+C_{22}A_{22}=E_{22}.$$

которая в условиях теоремы таким же образом разрешима по лемме. Итак, в условиях теоремы оператор A имеет правый и левый обратный; по I.I3в оператор A обратим, что и требуется.

Приведенные условия обратимости оператора A по его матрице являются лишь достаточными, но вовсе не необходимыми. Это иллюстрирует пример с 4×4-матрицей

$$A\cong\left\|\begin{array}{cc|cc}1&0&0&0\\0&0&1&0\\\hline 0&1&0&0\\0&0&0&1\end{array}\right\|$$

Здесь $\det A=-1$, так что A обратимая матрица; однако, ни один из четырех операторов $R_2\to R_2$, опреде-

ляемых выделенными минорами второго порядка, не является обратимым.

I.15. Непрерывные функции.

а. Чтобы иметь возможность говорить о непрерывности функции $f(x)$ $(M \to Y)$, мы должны предположить, что множество M, где определена эта функция, и множество Y, в котором она принимает свои значения, суть метрические пространства. Напомним, что множество M называется метрическим пространством, если на нем определена метрика-числовая функция двух точек $\rho(x,y)$ (или, при необходимости указать пространство, $\rho_M(x,y)$), удовлетворяющая условиям:

$\rho(x,y)=\rho(y,x)$ для любых x и y из M;

$\rho(x,z) \leq \rho(x,y)+\rho(y,z)$ для любых x,y,z из M;

$\rho(x,x)=0$ для любого $x \in M$, $\rho(x,y)>0$ при $x \neq y$.

Последовательность $x_1,\dots,x_n$ точек M называется сходящейся к точке $a \in M$, если $\lim\limits_{n\to\infty} \rho(x_n,a)=0$. Последовательность $x_1,\dots,x_n,\dots$ точек M называется фундаментальной, если $\lim\limits_{n,m\to\infty} \rho(x_n,x_m)=0$. Пространство M называется полным, если всякая фундаментальная последовательность является сходящейся, компактным, если всякая последовательность содержит фундаментальную подпоследовательность, и компактом, если оно полно и компактно. (Вещественная ось с метрикой $\rho(a,b)=|a-b|$ есть полное, но не компактное пространство; интервал на оси - компактное, ноне полное; отрезок - компакт). Мы предполагаем известными понятия замкнутого множества,открытого множества, сферы и замкнутого и открытого шара.

Две метрики ρ и ρ' на одном и том же множестве M называются эквивалентными, если они порождают одну и ту же сходимость, иначе говоря, если из $\rho(x_n,a)\to 0$ следует $\rho'(x_n,a)\to 0$ и обратно. (Так, например, любая метрика $\rho(x_1,x_2)$ эквивалентна метрике $\rho'(x_1,x_2)= = \operatorname{arctg} \rho(x_1,x_2)$).

б. Функция $y=f(x)$ называется непрерывной при $x=a$, если для каждого $\varepsilon>0$ существует такое $\delta>0$, что из $\rho_M(x,a)<\delta$, $x\in M$ следует $\rho_Y(f(x),f(a))<\varepsilon$.

Имеется и второе, эквивалентное определение непрерывности: функция $f(x)$ непрерывна при $x=a$, если для любой последовательности $x_1, \ldots, x_m, \ldots$ точек множества M, сходящейся к точке a, имеет место соотношение $f(x_m) \to f(a)$. Доказательство эквивалентности этих определений общеизвестно и мы не будем на нем здесь останавливаться.

Поскольку при определении непрерывности функции $f(x): M \to Y$ здесь используется лишь понятие сходимости $x_m \to a$, эквивалентные метрики ρ' и ρ на пространстве M приводят к одному и тому же запасу непрерывных функций.

Функция $f(x): M \to Y$, непрерывная в каждой точке множества M называется непрерывной на множестве M или $C^0(M, Y)$ - функцией, или просто - если M и Y известны из контекста - C^0- функцией.

в. Примеры. Самым простым примером непрерывной функции является постоянная - функция $f(x)$ $(M \to Y)$, которая при всех $x \in M$ принимает одно и то же значение $y_0 \in Y$.

Другим простым примером является числовая функция $f(x) = \rho(x,a)$ $(M \to R_1)$, где a - фиксированная точка пространства M. Ее непрерывность следует из неравенства треугольника.

Функция $f(x) \equiv x$ $(M \to M)$, ставящая в соответствие каждому элементу x метрического пространства M сам этот элемент x, является простейшим примером непостоянной непрерывной функции.

Для C^0- функции $y = f(x): M \to Y$ полный прообраз всякого открытого (замкнутого) в Y множества является открытым (замкнутым) в M. В частности, поверхность уровня непрерывной функции $f(x)$, т.е. множество тех $x \in M$, для которых $f(x) = y_0$ (фиксированный элемент из Y) является замкнутым множеством в X.

г. Прямое произведение метрических пространств. Пусть $X_1, \ldots, X_n$ - метрические пространства. Образуем их прямое произведение $X = X_1 \times \cdots \times X_n$ (I.IIв) и

введем в него какую-либо способом метрику так, чтобы сходимость по ней была равносильна сходимости по координатам (т.е. чтобы сходимость последовательности $x^m = \{x_1^m, \ldots, x_n^m\} \in X$ к пределу $a = \{a_1, \ldots, a_n\} \in X$ имела место тогда и только тогда, когда $x_1^m \to a_1$ в $X_1, \ldots,$ $x_n^m \to a_n$ в X_n). Метрическое пространство X с этими свойствами будем называть _метрическим прямым произведением пространств_ $X_1, \ldots, X_n$. Подходящую метрику в X можно задать, например, по формуле

$$\rho_X(\{x_1, \ldots, x_n\}, \{x_1', \ldots, x_n'\}) = \max\{\rho_{X_1}(x_1, x_1'), \ldots, \rho_{X_n}(x_n, x_n')\} \quad (1)$$

Можно использовать и иные формулы, например,

$$\rho_X(\{x_1, \ldots, x_n\}, \{x_1', \ldots, x_n'\}) = \sum_{i=1}^{n} \rho_{X_i}(x_i, x_i') \quad (2)$$

или

$$\rho_X(\{x_1, \ldots, x_n\}, \{x_1', \ldots, x_n'\}) = \sqrt{\sum_{i=1}^{n} \rho_{X_i}^2(x_i, x_i')} \quad (3)$$

д. _Непрерывные функции нескольких переменных._ Пусть $X_1, \ldots, X_n$ и Y метрические пространства. Образуем метрическое прямое произведение $X = X_1 \times \cdots \times X_n$ (г) Пусть функция $f(x): X \to Y$ непрерывна в точке $x = a$ так, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что из $\rho_X(x, a) < \delta$ следует $\rho_Y(f(x), f(a)) < \varepsilon$. Так как определение непрерывности не зависит от выбора метрики (в классе всех метрик, эквивалентных данной) , то пусть при этом ρ_X выбрано по формуле (I), так что неравенство $\rho_X(x, a) < \delta$ равносильно системе неравенств $\rho_{X_i}(x_i, a_i) < \delta$. Поэтому непрерывность функции $f(x)$ в точке a можно определить и как "непрерывность при $x = a$ по совокупности аргументов $x_1, \ldots, x_n$": для любого $\varepsilon > 0$ существует такое $\delta > 0$, что из $\rho_{X_i}(x_i, a_i) < \delta \ (i = 1, \ldots, n)$ следует $\rho_Y(f(x), f(a)) < \varepsilon$.

I.16. Свойства непрерывных функций.

а. Имеет место важная теорема:

Теорема о непрерывности сложной функции: если $y=f(x)\,(M\to Y)$ – функция, определенная в метрическом пространстве M, принимающая значения в метрическом пространстве Y, непрерывная в точке $x=a$, а $g(y)\,(Y\to Z)$ – функция, определенная в метрическом пространстве Y, принимающая значения в метрическом пространстве Z и непрерывная при $y=f(a)$, то сложная функция $\phi(x) = g[f(x)]\ (M\to Z)$ (заведомо определенная в некоторой окрестности точки a) также непрерывна при $x=a$.

Доказательство см., напр., в 05.15.

б. Если функции $f_1(x),\dots,f_m(x)$ определены на одном и том же множестве M, принимают значения в метрическом пространстве Y и образуют на M сходящуюся последовательность, то можно определить на том же пространстве M предельную функцию $f(x)=\lim\limits_{m\to\infty} f_m(x): M\to Y$. Если M есть метрическое пространство, сходимость $f_m(x)$ к $f(x)$ равномерна и функции $f_m(x)$ $(m=1,2,\dots)$ непрерывны, то и $f(x)$ непрерывна (05.96).

в. Множество B в метрическом пространстве Y называется ограниченным, если числа $\rho(y',y'')$, $y'\in B$, $y''\in B$ ограничены в совокупности. Функция $y=f(x): M\to Y$ с значениями в метрическом пространстве Y называется ограниченной (на M), если множество ее значений ограничено. Совокупность $Y(M)$ всех непрерывных ограниченных функций, определенных в метрическом пространстве M и принимающих значения в метрическом пространстве Y, сама становится метрическим пространством при вводе метрики по формуле

$$\rho_{Y(M)}(f,g) = \sup_{x\in M} \rho_Y(f(x),g(x))$$

Сходимость в смысле метрики $\rho_{Y(M)}$ есть равномерная сходимость на M. Если Y полно, то и $Y(M)$ полно (012.23е).

г. Пусть M и Y метрические пространства; функция $y=f(x)\ (M\to Y)$ называется равномерно непрерывной (на M), если для любого $\varepsilon>0$ можно найти

такое $\delta > 0$, что из $\rho_M(x', x'') < \delta$ следует $\rho_Y(f(x'), f(x'')) < \varepsilon$. На компакте (I.I5 а) всякая непрерывная функция равномерно непрерывна (05.I7б). Множество всех значений непрерывной функции на компакте M всегда компактно в Y (05.I6а), и, в частности, всегда ограничено. Поэтому для компакта M пространство $Y(M)$ есть пространство всех непрерывных функций $\varphi(x): M \to Y$.

д. Следующий способ построения более сложных непрерывных функций из более простых будет встречаться в дальнейшем. Пусть M, Y и Z метрические пространства и пусть дана функция $z = \Phi(x, y): M \times Y \to Z$, ограниченная и равномерно непрерывная на $M \times Y$. Пусть далее имеется непрерывная функция $f(x)$ $(M \to Y)$; тогда функция $\Phi(x, f(x)): M \to Z$ ограничена и, в силу а, непрерывна. Таким образом, определено отображение $F(f): f(x) \to F(f(x)) = \Phi(x, f(x))$ пространства $Y(M)$ в пространство $Z(M)$. Покажем, что это отображение непрерывно. Действительно, для любого $\bar{f} = \bar{f}(x) \in Y(M)$ мы имеем

$$\rho_{Z(M)}(F(\bar{f}), F(f)) = \sup_{x \in M} \rho_Z(\Phi(x, \bar{f}(x)), \Phi(x, f(x))) \qquad (I)$$

Для заданного $\varepsilon > 0$, используя предположенную равномерную непрерывность функции $\Phi(x, y)$, можно найти такое $\delta > 0$, что при $\rho_{Y(M)}(\bar{f}, f) < \delta$, т.е. $\sup_{x \in M} \rho_Y(\bar{f}(x), f(x)) < \delta$ правая часть в (I) станет меньше, чем ε, что нам и требуется.

I.I7. Непрерывные линейные операторы.

а. Чтобы говорить о совместных свойствах линейности и непрерывности, необходимо иметь пространство, соединяющее в себе структуры линейного и метрического пространств.

Наиболее удобным представляется для использования линейное нормированное пространство , т.е. линейное пространство, снабженное нормой - функцией, ставящей в соответствие каждому $x \in X$ вещественное число $|x|$, подчиненное условиям:

(I) $|0| = 0$, $|x| > 0$ при $x \neq 0$;

(2) $|\alpha x| = |\alpha| \cdot |x|$ для любого вещественного α;

(3) $|x+y| \le |x| + |y|$ для любых x и y из X.

Нормированное линейное полное пространство называется банаховым.

Пусть X и Y нормированные линейные пространства; будем рассматривать линейные операторы $A: X \to Y$, являющиеся непрерывными функциями в той метрике, которая задается нормами в пространствах X и Y:

$$\rho_X(x_1, x_2) = |x_1 - x_2|, \quad \rho_Y(x_1, x_2) = |y_1 - y_2|$$

Такие функции будем называть непрерывными линейными операторами. Так линейные операторы в конечномерных нормированных пространствах (I.14б) всегда непрерывны, что видно непосредственно из их общего представления.

Композиция $C = BA$ непрерывных линейных операторов $A: X \to Y$ и $B: Y \to Z$ является также непрерывным линейным оператором $(X \to Y)$.

б. Непрерывные линейные операторы $A: X \to Y$ образуют линейное пространство (подпространство линейного пространства всех линейных операторов, действующих из X в Y). Мы в дальнейшем будем употреблять обозначения $L(X, Y)$ и $L(X)$ только для пространств линейных непрерывных операторов. Пространство $L(X, Y)$ можно также нормировать по формуле

$$\|A\| = \sup_{|x|=1} |Ax|$$

при этом для любого $x \in X$

$$|Ax| \le \|A\| \, |x|$$

С введением этой нормы пространство $L(X, Y)$ (в частности, $L(X)$) становится нормированным линейным пространством, причем полным, если полно Y.

в. Прямая сумма нормированных пространств. Пусть $X_1, \dots, X_n$ — нормированные линейные пространства. Образуем их прямую сумму $X = X_1 + \dots + X_n$ (I.14е) и введем в нее каким-либо образом норму так, чтобы сходимость по ней была равносильна сходимости по координатам, (как в прямом произведении метрических пространств I.15г). Пространство X с такой нормой будем называть нормированной прямой

суммой пространств $X_1, \ldots, X_n$, впрочем, слово "нормированная" будем, как правило, опускать. Подходящую норму в X можно задать, например, по формуле

$$|\{x_1, \ldots, x_n\}| = \max_{1 \le k \le n} |x_k|_k$$

где $|x_k|_k$ есть норма элемента x_k в пространстве X_k. При этом составляющие x_k естественно оказываются непрерывными функциями от вектора x, так что линейные операторы проектирования $\mathcal{P}_i$ ($\mathcal{P}_i x = x_i$) непрерывны.

Как и в I.14е, пространства X_k можно отождествить с некоторыми подпространствами $X'_k \subset X$. В дополнение к свойствам (α)-(β), отмеченным в I.14е, каждое подпространство X_k оказывается замкнутым в X (как совокупность общих решений уравнений $\mathcal{P}_i x = 0$, $i \neq k$ с непрерывными функциями $\mathcal{P}_i$).

Обратно, если в пространстве X выделены каким-то образом подпространства $X_1, \ldots, X_n$ так, что выполняются условия $(\alpha)-(\beta)$ из I.14е, т.е. для каждого $x \in X$ имеется разложение $x = x_1 + \cdots + x_n$ и это разложение для каждого x единственно, и если при этом составляющие $x_1, \ldots, x_n$ - непрерывные функции от x, мы будем говорить, что пространство X разложено в (нормированную) прямую сумму подпространств $X_1, \ldots, X_n$ (Можно доказать, что если X полно, а подпространства $X_1, \ldots, X_n$ замкнуты, то из условий I.14е уже вытекает непрерывность составляющих x_k).

Если пространство X разложено в прямую сумму подпространств $X_1, \ldots, X_n$, а пространство Y - в прямую сумму подпространств $Y_1, \ldots, Y_m$, и задан линейный непрерывный оператор $A: X \to Y$, то линейные операторы матричного представления $A_{ij} = Q_j A \mathcal{P}_i : X_i \to Y_j$ также являются непрерывными. Обратно, любому набору линейных непрерывных операторов A_{ij} $(i=1,\ldots,n, j=1,\ldots,m)$ соответствует по правилам I.14ж линейный непрерывный оператор $A: X \to Y$ с матрицей $\|A_{ij}\|$.

I.18. Произвольные непрерывные функции в нормированных пространствах.

а. Если $f(x)$ и $g(x)$ - две функции, определенные на

одном и том же множестве X и принимающие значения в линейном пространстве Y, то можно определить на том же множестве X функцию $y(x) = f(x) + g(x)$ $(X \to Y)$ сумму функций $f(x)$ и $g(x)$. Если X - метрическое пространство, а Y - нормированное линейное пространство и функции $f(x)$ и $g(x)$ непрерывны, то и $y(x)$ непрерывна

б. Если функция $y(x)$ определена на множестве X и принимает значения в линейном пространстве Y, а C - линейный оператор, отображающий линейное пространство Y в линейное пространство Z, то на том же множестве X можно определить функцию $z(x) = Cy(x)$. Если при этом X - метрическое пространство, Y и Z - нормированные пространства и функция $y(x)$ и оператор C непрерывны, то и функция $z(x)$ непрерывна (I.I6а).

в. Можно рассматривать разного вида произведения функций $\lambda(x)$ и $f(x)$, определенных на одном и том же множестве X. Например, такое произведение определено, если $f(x)$ отображает X в линейное пространство Y, а $\lambda(x)$ - числовая функция. Результатом является функция $(\lambda f)(x) = \lambda(x) \cdot f(x) (X \to Y)$. Оно определено также, если $\lambda(x)$ и $f(x)$ принимают значения в одной и той же алгебре X и его значения принадлежат той же алгебре X. Оно определено также для $f(x): X \to Y$ и для $\lambda(x): X \to L(Y, Z)$, и тогда функция $\lambda(x) \cdot f(x)$ отображает X в Z. Общим свойством всех этих произведений является билинейность результата по каждому из множителей.

Введем следующее общее определение. Пусть на нормированной прямой сумме нормированных пространств Λ и F определен непрерывный оператор $B(\lambda, f)$ $(\lambda \in \Lambda, f \in F)$, линейный по каждому из аргументов λ и f и принимающий значения в нормированном пространстве B; будем называть его обобщенным произведением элементов λ и f и обозначать $b = \langle \lambda, f \rangle$. Пусть, далее, имеются функции $\lambda(x): X \to \Lambda$ и $f(x): X \to F$, тогда определена функция $\langle \lambda(x), f(x) \rangle$ $(X \to B)$, которая называется обобщенным произведением функций $\lambda(x)$ и $f(x)$. Три приведенных выше конкретных произведения входят в эту схему, если положить, соответственно,

$\langle \lambda(x), f(x) \rangle = \lambda(x) \cdot f(x)$ в обычном смысле произведения $f(x)$ на число $\lambda(x)$;

$\langle \lambda(x), f(x)\rangle = \lambda(x)\cdot f(x)$ в смысле произведения элементов алгебры;

$\langle \lambda(x), f(x)\rangle = \lambda(x)\cdot f(x)$ в смысле применения оператора $\lambda(x)$ к вектору $f(x)$.

Утверждается, что обобщенное произведение $b(x) = \langle \lambda(x), f(x)\rangle$ есть непрерывная функция от x, если непрерывны $\lambda(x)$ и $f(x)$. Действительно, $b(x)$ есть композиция двух функций:

$$x \to \{\lambda(x), f(x)\} \in \Lambda + F \text{ и } \{\lambda, f\} \to \langle \lambda, f\rangle \in B$$

Первая из них непрерывна согласно предположению о непрерывности $\lambda(x)$ и $f(x)$; и по определению нормированной прямой суммы; вторая - по предположению непрерывности функции $\langle \lambda, f\rangle$: $\Lambda + F \to B$. По теореме о непрерывности сложной функции непрерывен и результат $b(x) = \langle \lambda(x), f(x)\rangle$, что и утверждалось.

В частности, являются непрерывными и три рассмотренных выше конкретных произведения при условии, что непрерывен каждый из множителей.

г. Результаты а и в справедливы, в частности, для числовых функций $X \to R_1$, непрерывных на метрическом пространстве X.

В частности, так как в пространстве $X = R_n$ координаты точки $x = \{x_1, \dots, x_n\}$ являются непрерывными функциями от x и функция $1/t$ $(R_1 \to R_1)$ непрерывна при $t \neq 0$, то по а, в и I.16а непрерывными являются многочлены и рациональные функции от координат (последние - вне нулей знаменателя). По I.16б предел последовательности многочленов, равномерно сходящейся на некотором множестве $G \subset R_n$, есть непрерывная функция $(G \subset R_1)$ на G. Для некоторых множеств $G \subset R_n$ справедливо и обратное: каждая непрерывная функция $f(x)$ $(G \to R_1)$ может быть представлена как предел равномерно сходящейся последовательности многочленов; это справедливо, в частности, для компактов (012.44).

§ I.2. Дифференцируемые функции

Основная идея дифференциального исчисления состоит в замене данной функции в окрестности некоторой точки линейной функцией с ошибкой более высокого порядка малости, чем соответствующее

приращение аргумента. Те числовые функции одного переменного x, для которых такая замена возможна, составляют класс дифференцируемых функций от x. Но возможность локальной аппроксимации линейной функцией вовсе не требует одномерности аргумента. Сначала мы приведем необходимые определения для числовой функции нескольких переменных, а затем дадим общее определение, пригодное для общего случая, когда и область определения, и область значений функции являются нормированными линейными пространствами.

I.21. Прежде всего напомним основное определение дифференцируемой числовой функции вещественного переменного.

Числовую функцию $f(x)$ вещественного переменного x, $a \leq x \leq b$ мы называли дифференцируемой в точке $c \in (a,b)$, если существовал предел

$$\lim_{h\to 0} \frac{f(c+h) - f(c)}{h} = f'(c) \qquad (1)$$

В этом случае в приращении функции $f(x)$ при переходе аргумента от значения $x=c$ к значению $x=c+h$ можно выделить главную часть, которая зависит от h линейно:

$$f(c+h) - f(c) = f'(c)\,h + o(h), \quad \lim_{h\to 0} \frac{o(h)}{h} = 0$$

Обратно, если приращение функции $f(x)$ при переходе от значения $x=c$ к значению $x=c+h$ допускает выделение главной части, линейной по h, т.е. при некоторой постоянной D выполняется соотношение

$$f(c+h) - f(c) = Dh + o(h), \quad \lim_{h\to 0} \frac{o(h)}{h} = 0, \qquad (2)$$

то предел $\lim\limits_{h\to 0} \dfrac{f(c+h) - f(c)}{h}$ существует и равен D.

Таким образом, для числовых функций одного переменного соотношения (1) и (2) служат эквивалентными определениями дифференцируемости в точке $x=c$.

I.22. а. Теперь рассмотрим случай числовой функции $f(x)$ от n вещественных переменных $x=(x_1,\dots,x_n)$. Из двух приведенных выше определений на этот случай наиболее естественно переносится второе.

Именно, функцию $f(x) = f(x_1,\dots,x_n)$, $f: R_n \to R_1$,

мы будем называть дифференцируемой в точке $x = c = (c_1, \ldots, c_n)$, если приращение $f(x)$ при переходе из точки $x = c$ в точку $x = c + h$, $h = (h_1, \ldots, h_n)$ допускает выделение главной части, линейной по h. Последнее означает, что при некоторых постоянных $\mathcal{D}_1, \ldots, \mathcal{D}_n$ выполняется соотношение

$$f(c+h) - f(c) = \sum_{i=1}^{n} \mathcal{D}_i h_i + o(h), \quad \lim_{h \to 0} \frac{o(h)}{|h|} = 0. \tag{I}$$

Величина $\sum_{i=1}^{n} \mathcal{D}_i h_i$ и есть главная линейная часть приращения функции $f(x)$ при изменении x от c до $c+h$. Величина $o(h)$ есть малая высшего порядка сравнительно с h. При этом $\lim_{h \to 0} \frac{o(h)}{|h|} = 0$ точно означает следующее: для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что из $|h| < \delta$ следует $\frac{|o(h)|}{|h|} < \varepsilon$.

Под $|h|$ здесь понимается любая норма в пространстве R_n, например, $\sqrt{\sum_{i=1}^{n} h_i^2}$. Поскольку все нормы в R_n эквивалентны (012.35д), если соотношение $\lim_{h \to 0} \frac{o(h)}{|h|} = 0$ выполняется для одной какой-либо нормы $|h|$, то оно выполняется и для любой другой нормы.

Определение (I) по форме зависит от выбора базиса и системы координат в пространстве R_n. Но если вспомнить, что линейная форма $\sum_{i=1}^{n} \mathcal{D}_i h_i$ при переходе к новому базису остается линейной формой (с новыми коэффициентами), так что в новом базисе слагаемые в разложении (I) сохранят свои значения, изменив лишь форму записи – то станет ясно, что дифференцируемость функции $f(x)$ при $x = c$ есть ее внутреннее свойство, не зависящее от выбора базиса в пространстве R_n.

б. Если же базис и тем самым система координат фиксированы, то можно утверждать, что для дифференцируемой функции $f(x)$ коэффициенты $\mathcal{D}_i$ в формуле (I) определяются единственным образом. Чтобы убедиться в этом, возьмем какое-либо целое m между 1 и n и положим в формуле (I) $h = (0, \ldots, 0, h_m, 0, \ldots, 0)$,

Тогда мы получим $|h| = |h_m|$ и

$$f(c+h)-f(c) = f(c_1,\ldots,c_{m-1},c_m+h_m,c_{m+1},\ldots,c_n)-$$
$$-f(c_1,\ldots,c_{m-1},c_m,c_{m+1},\ldots,c_n) = D_m h_m + o(h_m);$$

это означает, что функция $f(c_1,\ldots,x_m,\ldots,c_n)$ одного переменного x_m дифференцируема по этому переменному x_m в точке $x_m = c_m$ и число D_m есть как раз ее производная по этому аргументу:

$$D_m = \lim_{h_m\to 0}\frac{f(c_1,\ldots,c_m+h_m,\ldots,c_n)-f(c_1,\ldots,c_m,\ldots,c_n)}{h_m} \qquad (2).$$

Полученная явная формула для D_m доказывает единственность коэффициентов в формуле (I).

Если предел в правой части (2) существует, то он называется частной производной от функции $f(x)$ в точке $x=c$ по переменному x_m. Таким образом, у функции $f(x)$, дифференцируемой (в смысле (I)) в точке $x=c$, имеется в этой точке частная производная по любому из переменных $x_1,\ldots,x_n$.

Частная производная функции $f(x)$ по переменному x_m в точке $x=c$ обозначается через $\frac{\partial f(c)}{\partial x_m}$ или $f'_{x_m}(c)$. Заметим, что частные производные по всем переменным у функции $f(x)$ могут существовать в точке c и без того, чтобы функция $f(x)$ была дифференцируема в точке c (см.задачу 24).

в. Из чисел $D_1,\ldots,D_n$ можно построить вектор $D = \{D_1,\ldots,D_n\}$. Он называется градиентом функции $f(x)$ в точке $x=c$ и обозначается через $\operatorname{grad} f(c)$. Выражение $\sum_{i=1}^{n} D_i h_i$ называется дифференциалом функции $f(x)$ в точке c, отвечающим смещению $h=\{h_1,\ldots,h_n\}$. Оно обозначается также через $df(c)$. Величины h и h_i традиционно обозначаются через dx и dx_i $(i=1,\ldots,n)$ соответственно. Если в пространстве R_n ввести скалярное произведение векторов $x=\{x_1,\ldots,x_n\}$ и $y=\{y_1,\ldots,y_n\}$ по формуле $(x,y)=\sum_{i=1}^{n} x_i y_i$,

то дифференциал функции $f(x)$ в точке c можно записать в любой из форм

$$df(c)=\sum_{i=1}^{n}D_i h_i=(D,h)=(grad\, f(c),h)=\sum_{i=1}^{n}\frac{\partial f(c)}{\partial x_i}dx_i \quad (3)$$

Формула (I) приобретает при этом вид

$$f(c+h)-f(c)=df(c)+o(h)=(grad\, f(c),h)+o(h)=$$

$$=\sum_{i=1}^{n}\frac{\partial f(c)}{\partial x_i}dx_i+o(dx) \quad (4)$$

г. Формула (4) может служить для сравнительной оценки изменения функции при смещении аргумента из точки c по различным направлениям. Именно, если w есть угол между векторами $grad\, f(c)$ и h, то по определению скалярного произведения имеем

$$f(c+h)-f(c)=(grad\, f(c),h)=|grad\, f(c)|\cdot|h|\cdot\cos w$$

Отсюда при $grad\, f(c)\neq 0$ можно сделать следующие выводы (рис. I.2-I).

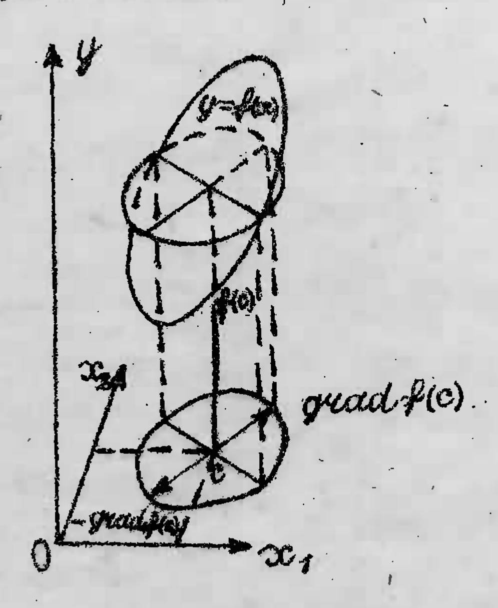


Рис. I.2I

При смещении в направлении вектора $grad\, f(c)$ $(\cos w=1)$ функция $f(x)$ получает наибольшее возможное приращение на каждую единицу длины вдоль вектора смещения h (с точностью до малых высшего порядка). При смещении в противоположном направлении $(\cos w=-1)$ функция $f(x)$ получает отрицательное приращение по модулю опять таки наибольшее возможное на каждую единицу длины вдоль вектора смещения h. Эти направления, связанные с вектором градиента, называются, естественно, направлением быстрейшего подъема и

направлением быстрейшего спуска функции $f(x)$ из точки $x=c$. При смещении в любом направлении, ортогональном вектору градиента $(\cos\omega=0)$, приращение функции $f(x)$ имеет порядок малости более высокий, чем $|h|$. Таким образом, вектор градиента указывает и направление быстрейшего подъема функции $f(x)$, и характеризует его величину.

д. Для функции двух переменных $y=f(x_1,x_2)\,(R_2\to R_1)$ более традиционно обозначение независимых переменных через x, y, а самой функции – через z,

$$z=f(x,y)$$

В этих обозначениях формулы (3) и (4) приобретают вид

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy \qquad (5)$$

$$f(x+dx,y+dy)-f(x,y)=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy+o(|dx|+|dy|) \qquad (6)$$

I.23. а. Теперь сформулируем общее определение дифференцируемой функции. Пусть функция $y=f(x)$ определена в некоторой области $G\subset X$ нормированного пространства X и принимает значения в нормированном пространстве Y. Эта функция называется дифференцируемой в точке $x=c\in G$, если приращение функции $f(x)$ при переходе из точки c в точку $c+h\in G$ допускает выделение главной части, линейной по h, т.е. если выполнено соотношение

$$f(c+h)-f(c)=Dh+o(h) \qquad (I)$$

где D есть некоторый непрерывный линейный оператор, действующий из пространства X в пространство Y, а $o(h)$ есть вектор в пространстве Y, удовлетворяющий условию

$$\lim_{h\to 0}\frac{o(h)}{|h|}=0$$

Таким образом, выражение Dh в данном случае представляет собой главную линейную часть приращения функции $f(x)$ при изменении x от c до $c+h$.

Равенству (I) можно дать еще следующую геометрическую трактовку. Функция $f(x)$ осуществляет отображение области $G\subset X$ в пространство Y; если перенести начало коор-

динат пространства X в точку c, а начало координат пространства Y в точку $f(c)$, т.е. считать независимым переменным вектор $h = x - c$, а функцией - вектор $k = y - f(c)$, то получающееся отображение $h \to k$ аппроксимируется линейным отображением $k = Dh$ (с точностью до бесконечно малого члена $o(h)$ высшего порядка сравнительно с h). Можно сказать, что само отображение $y = f(x)$ вблизи точки $x = c$ допускает выделение главной линейной части.

б. Покажем, что оператор D, фигурирующий в формуле (I), определен однозначно. Допустим, что наряду с равенством (I) существует другое, аналогичное, представление разности $f(c+h) - f(c)$, именно:

$$f(c+h) - f(c) = D_1 h + o_1(h), \quad \lim_{h\to 0} \frac{o_1(h)}{|h|} = 0 \qquad (2)$$

Вычитая (2) из (I) и обозначая $D_2 = D - D_1$, мы найдем, что

$$D_2 h = o_2(h), \quad \lim_{h\to 0} \frac{o_2(h)}{|h|} = 0$$

Для заданного $\varepsilon > 0$ выберем $\delta > 0$ так, чтобы из $|h| \le \delta$ следовало $|o_2(h)| \le \varepsilon |h|$. Тогда на основании I.17.б мы получим

$$\|D_2\| = \sup_{|h|\le\delta} \frac{|D_2 h|}{|h|} = \sup_{|h|\le\delta} \frac{|o_2(h)|}{|h|} \le \varepsilon$$

Так как $\varepsilon > 0$ произвольно, то $\|D_2\| = 0$, откуда $D_2 = D - D_1 = 0$ и, следовательно, $D = D_1$ что и требовалось.

в. Оператор D в формуле (I), определенный, как мы показали, однозначно, называется производной функции $f(x)$ при $x = c$ и обозначается через $f'(c)$. Выражение Dh, т.е. вектор в пространстве Y, полученный применением оператора производной к вектору смещения h, называется дифференциалом функции $f(x)$ в точке c, отвечающим смещению h. Пишут, по аналогии с числовыми функциями вещественного переменного,

$$df(c) = Dh = f'(c)\,h = f'(c)\,dx$$

где $dx = h$ означает любой вектор пространства X.

Функция $f(x)$, дифференцируемая во всех точках области G, называется дифференцируемой в области G. Ее производная $f'(x)$ есть линейный оператор, действующий из X в Y и зависящий от точки $x \in G$. Ее дифференциал $d f(x) = f'(x) h$ есть функция двух переменных – вектора $h \in X$ и точки $x \in G$.

Переход от функции $f(x)$ к ее производной $f'(x)$ или к ее дифференциалу $f'(x) dx$ называется (как и в случае одного переменного) дифференцированием функции $f(x)$.

I.24 а. Итак, производная функции $f(x): G \subset X \to Y$ при $x = c$ есть некоторый линейный оператор, действующий из X в Y, который применяется к вектору смещения $h \in X$. Для числовой функции вещественного переменного $(G \subset R_1 \to R_1)$ в классическом анализе производная определялась как число, а не как оператор. Это и не удивительно, поскольку всякий линейный оператор из R_1 в R_1 есть оператор умножения на некоторое число, и поэтому он может быть отождествлен с самим этим числом.

б. Рассмотрим теперь функцию $f(x): G \subset R_1 \to R_m$. Ее дифференцируемость при $x = c$ означает существование линейного оператора $\mathcal{D}: R_1 \to R_m$ такого, что имеет место соотношение

$$f(c+h) - f(c) = \mathcal{D} h + o(h) \qquad (1)$$

Здесь оператор $\mathcal{D}$ приводит в соответствие числовому смещению h вектор $\mathcal{D} h \subset R_m$, главную линейную часть приращения функции $f(x)$; этот оператор полностью определяется вектором $\mathcal{D} \cdot 1$ и может быть отождествлен с этим вектором. Вектор $\mathcal{D}$ может быть определен из (1) делением на h и переходом к пределу при $h \to 0$:

$$\mathcal{D} = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \qquad (2)$$

откуда возникает известная геометрическая интерпретация вектора $\mathcal{D}$ как вектора касательной к годографу (1.12б) функции $f(x)$ (рис. I.2-2). Тем не менее всее же производная есть не просто вектор определяемый правой частью (2), а оператор $R_1 \to R_n$, т.е. правило, сопоставляющее каждому числу $h \in R_1$ вектор $\mathcal{D} h \in R_n$.

$\mathcal{D} = f'(c)$ $\quad f(c+h)$ $\quad f(c)$

Рис. I.2-2

Заметим еще, что все сказанное здесь равно относится и к функции $y=f(x): G\subset X\Rightarrow R_1\to Y$, где Y -любое нормированное пространство.

в. Другое дело, когда $X\neq R_1$. Пусть, например, $f(x): G\subset R_n\to R_1$. Тогда, как мы знаем, дифференциал функции $f(x)$ при $x=c$ имеет вид

$$df(c)=\sum_{i=1}^{n}\mathcal{D}_i h_i \qquad (3)$$

Снова с производной можно связать вектор $\mathcal{D}=\{\mathcal{D}_1,\dots,\mathcal{D}_n\}\subset R_n$ - градиент функции $f(x)$ при $x=c$. Но этот вектор, как оператор, действует уже не так, как вектор $f'(c)$ в предыдущем случае $(R_1\to R_n)$. Правило действия градиента на вектор смещения $h=(h_1,\dots,h_n)$ задается теперь формулой (3), отличной от формулы (I). Это подчеркивает важность задания производной, как оператора, т.е. как правила действия на вектор смещения, каковы бы ни были соблазны связывать производную (по числу параметров и т.п.) с тем или иным геометрическим объектом.

г. Если $f(x): G\subset X\to R_1$ есть числовая функция, производная $f'(c)$ есть линейный оператор из X в R_1, т.е. линейный непрерывный функционал. В случае $X=R_n$ этот линейный функционал имеет вид (3). Имеются и другие пространства, где известный общий вид линейного функционала позволяет записывать производную аналогичным образом. Так, для случая, когда X есть полное гильбертово пространство H, известно, что каждый непрерывный линейный функционал $\mathcal{D}(h)$ имеет вид скалярного произведения $(\mathcal{D},h)$, где $\mathcal{D}$ есть (однозначно- определенный) вектор того же пространства H. Поэтому все построения I.22 г. могут быть полностью перенесены с R_n на H. И здесь также с производной $f'(c)$ можно сопоставить вектор $\mathcal{D}=grad\,f(c)$, действие которого на вектор смещения h задается скалярным произведением $(\mathcal{D},h)$ и направление которого есть направление быстрейшего возрастания функции $f(x)$.

д. Если X есть нормированное, но не гильбертово пространство (а Y - попрежнему есть R_1), то оператор $f'(c): X\to R_1$, вообще говоря, уже не задается вектором из X; он есть элемент сопряженного пространства X', всех непрерывных линейных функционалов на X, и действие

$f'(c)$ есть действие соответствующего линейного функционала. Уже не существует, вообще говоря, однозначного определенного вектора смещения h, обеспечивающего быстрейший рост функции $f(x)$ из точки $x=c$; может не быть ни одного такого вектора, или их может существовать много. (Задача 34).

I.25. а. Рассмотрим более детально векторные дифференцируемые функции в конечномерном случае.

Пусть функция $f(x): G\subset R_n \to R_m$ определена в области G n-мерного пространства $X=R_n$ и принимает значения в m-мерном пространстве $Y=R_m$. Выбрав в обоих этих пространствах каким-либо образом базисы и записывая $x=\{x_1,\ldots,x_n\}\in R_n$, $y=\{y_1,\ldots,y_m\}\in R_m$ мы можем выразить векторную функцию $y=f(x)$ с помощью m числовых функций

$$\left.\begin{aligned} y_1 &= f_1(x) = f_1(x_1,\ldots,x_n) \\ &\ldots\ldots\ldots \\ y_m &= f_m(x) = f_m(x_1,\ldots,x_n) \end{aligned}\right\} \qquad (1)$$

Пусть функция $f(x)$ дифференцируема при $x=c$. Равенство

$$f(c+h)-f(c) = f'(c)\,h + o(h) \qquad (2)$$

определяет линейный оператор $f'(c): R_n \to R_m$. Как и всякому линейному оператору, действующему из R_n в R_m оператору $f'(c)$ можно поставить в соответствие некоторую $n\times m$-матрицу. Для этого нужно равенство (2) записать в координатной форме, используя имеющийся базис в пространстве R_n и в пространстве R_m; при этом мы получаем

$$f_i(c+h) - f_i(c) = \sum_{j=1}^{n} f'_{ij}(c)h_j + o_i(h) \quad (i=1,\ldots,m) \qquad (3)$$

Величины $f'_{ij}(c)$ $(j=1,\ldots,n;\ i=1,\ldots,m)$ и составляют $n\times m$-матрицу оператора $f'(c)$ относительно указанных базисов. Формулы (3) показывают, что составляющие $f_i(x)$ дифференцируемой (при $x=c$) векторной функции $f(x)$ сами являются дифференцируемыми (при $x=c$) числовыми функциями. Очевидно, что справедливо и обратное: если числовые функции $f_i(x)$, $i=1,\ldots,m$, дифференцируемы при $x=c$, то и векторная функция $f(x)$ также дифференцируема при $x=c$. Мы видим также, каковы элементы матрицы линейного

оператора $f'(c)$: поскольку коэффициенты главной линейной части приращения любой числовой функции суть ее частные производные по соответствующим переменным, мы имеем

$$f'_{ij}(c) = \frac{\partial f_i(c)}{\partial x_j} \quad (i=1,\dots,m;\ j=1,\dots,n) \tag{4}$$

Итак, матрица линейного оператора $f'(c)$ $(R_n \to R_m)$ составлена из частных производных и имеет вид

$$f'(c) \cong \left\| \begin{matrix} \frac{\partial f_1(c)}{\partial x_1} & \frac{\partial f_1(c)}{\partial x_2} & \dots & \frac{\partial f_1(c)}{\partial x_n} \\ \frac{\partial f_2(c)}{\partial x_1} & \frac{\partial f_2(c)}{\partial x_2} & \dots & \frac{\partial f_2(c)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m(c)}{\partial x_1} & \frac{\partial f_m(c)}{\partial x_2} & \dots & \frac{\partial f_m(c)}{\partial x_n} \end{matrix} \right\|$$

Здесь знак $\cong$ означает: оператору $f'(c)$ при выбранных базисах (координатах) в R_n и R_m отвечает матрица, выписанная справа. Эта матрица называется матрицей Якоби. Она обозначается также через $\left\| \frac{\partial f_i(c)}{\partial x_j} \right\|$ или $\left\| \frac{\partial (f_1,\dots,f_m)}{\partial (x_1,\dots,x_n)} \right\|$. В случае $n=m=1$ она состоит из одного элемента, совпадающего с обычной производной вещественной функции по вещественному переменному. Для числовой функции n вещественных переменных $m=1$ и матрица Якоби имеет одну строку (I.24.в.). Для m функций одного вещественного переменного (I.24.б.) $n=1$ и матрица Якоби превращается в одностолбцовую матрицу.

В случае $m=n$ матрица Якоби – квадратная матрица:

$$f'(c) \cong \left\| \begin{matrix} \frac{\partial f_1(c)}{\partial x_1} & \dots & \frac{\partial f_1(c)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(c)}{\partial x_1} & \dots & \frac{\partial f_n(c)}{\partial x_n} \end{matrix} \right\|$$

Ее определитель, как мы увидим далее, является важной характеристикой отображения $y=f(x)$ при $x=c$; он называется якобианом отображения $f(x)$ при $x=c$ и обозначается через

$$\det \left\| \frac{\partial (f_1,\dots,f_n)}{\partial (x_1,\dots,x_n)} \right\| \text{ или } \frac{\partial (f_1,\dots,f_n)}{\partial (x_1,\dots,x_n)}$$

Подчеркнем, что матрица Якоби только соответствует оператору $f'(c)$; действие же оператора $f'(c)$ с помощью соответствующей ему матрицы Якоби выражается равенством (3).

6. Особенно просто дело обстоит в случае линейной функции $y=f(x)$, или в развернутой форме,

$$\begin{aligned} y_1 &= D_{11}x_1+\dots+D_{1n}x_n, \\ &\dots\dots\dots\dots \\ y_m &= D_{m1}x_1+\dots+D_{mn}x_n \end{aligned}$$

где D_{ij} суть постоянные числа. Матрица Якоби в этом случае представляет собой матрицу, составленную из чисел D_{ij}:

$$f'(c) \cong \begin{Vmatrix} D_{11} & \dots & D_{1n} \\ \dots & \dots & \dots \\ D_{m1} & \dots & D_{mn} \end{Vmatrix}$$

и не зависит от точки c.

I.26. а. Для функции $y=f(x): G\subset R_1\to R_1$ (вещественной функции вещественного переменного), как мы знаем существование производной числовой функции $y=f(x)$ в точке $x=c$ с геометрической точки зрения означает существование в точке $(c,f(c))$ касательной к кривой – графику функции $f(x)$ в плоскости R_2. При этом касательная определялась или как предельное положение секущей – что соответствует аналитическому определению I.21 (I), – или же как прямая, точки которой отстоят от соответствующих (т.е. взятых при тех же x) точек кривой $y=f(x)$ на бесконечно малую высшего порядка по сравнению с $h=x-c$; последнее соответствует аналитическому определению I.21 (2).

Угловой коэффициент касательной совпадает с производной $f'(c)$, а полное уравнение касательной имеет вид

$$y-p=f'(c)\,(x-c) \quad (p=f(c)) \qquad (I)$$

или же, обозначая $x-c=dx,\ y-p=dy,$

$$dy=f'(c)\,dx \qquad (2)$$

В случае функции $y=f(x): G\subset R_1\to Y$, где Y любое нормированное пространство, график ее представляет собою кривую в пространстве $R_1\times Y$. Уравнение касательной имеет попрежнему вид (I) (или (2)), с той разницей, что теперь $y-p$ и $f'(c)$ являются векторами пространства Y.

б. Для случая числовой функции $y=f(x_1,\dots,x_n)$ геометрическое содержание понятия дифференцируемости связано с наличием касательной плоскости к поверхности $y=f(x_1,\dots,x_n)$. Чтобы притти к определению касательной плоскости, будем обобщать второе из приведенных определений касательной прямой. Именно, плоскость $y-p=\sum_{i=1}^{n}A_i(x_i-c_i)$, $p=f(c)$, в $(n+1)$-мерном пространстве $x_1,\dots,x_n,y$ будем называть касательной плоскостью к поверхности $y=f(x)$ в точке $x=c$, если отклонение точек $\{x_1,\dots,x_n,f(x)\}$ и $\{x_1,\dots,x_n,p+\sum_{i=1}^{n}A_i(x_i-c_i)\}$ (первая из них на поверхности, вторая на плоскости) имеет более высокий порядок малости по сравнению с $|h|=\sqrt{\sum_{i=1}^{n}(x_i-c_i)^2}$. Ясно, что аналитическое условие I.22 (I) дифференцируемости функции $f(x)$ при $x=c$ теперь можно трактовать как условие существования при $x=c$ касательной плоскости к поверхности $y=f(x)$; в обозначениях I.22 (I) эта касательная плоскость имеет уравнение

$$y-p=\sum_{i=1}^{n}\mathcal{D}_i(x_i-c_i)=\sum_{i=1}^{n}\frac{\partial f(c)}{\partial x_i}(x_i-c_i) \qquad (3)$$

или, в дифференциалах $(dy=y-p,\ dx_i=x_i-c_i,\ i=1,\dots,n)$

$$dy=\sum_{i=1}^{n}\frac{\partial f(c)}{\partial x_i}dx_i \qquad (4)$$

в. Определение касательной плоскости можно дать и для самой общей функции $y=f(x): G\subset X\to Y$, дифференцируемой в точке $x=c$. Пусть $f(c)=p$. Тогда "касательной плоскостью" к поверхности $y=f(x)$ в точке c мы будем называть линейное многообразие в пространстве $X+Y$ (прямая сумма), выделенное уравнением

$$y - p = f'(c)(x-c) \quad (x \in \mathcal{B}, y \in \mathcal{Y}) \qquad (5)$$

И здесь отклонение точки $\{x, f(x)\}$ от точки $\{x, p + f'(c)(x-c)\}$ (первая – на поверхности $y = f(x)$, вторая – на многоообразии (5)) имеет более высокий порядок малости по сравнению с $|x-c|$. Используя дифференциалы, мы можем записать это уравнение в форме

$$dy = f'(c)\,dx \qquad (6)$$

§ I.3. Общие теоремы о дифференцируемых функциях.

I.3I. В дальнейшем функции $f(x), g(x), \ldots$ предполагаются действующими из некоторой окрестности V точки c нормированного линейного пространства X в нормированное линейное пространство $\mathcal{Y}$.

а. Если $f(x) = const$ (т.е. для всех $x \in V$ значение функции $f(x)$ есть один и тот же элемент пространства $\mathcal{Y}$), то $f'(c) = 0$. Этот результат вытекает из равенства $f(c+h) - f(c) = 0$ и единственности производной (I.23).

б. Если $f(x)$ есть линейная функция $F \cdot x$ (точнее, если существует линейный оператор $F: X \to \mathcal{Y}$, значения которого совпадают на $V(c)$ со значениями функции $f(x)$, то $f'(c) = F$, $df(c) = F\,dx$.

Действительно, в данном случае

$$f(c+h) - f(c) = F(c+h) - F(c) = F(h)$$

и требуемый результат также вытекает из единственности производной (I.23).

в. В частности, для функции $f(x) = x$ $(X \to X)$ производная $f'(x)$ есть тождественный оператор $f'(x) = E\,(X \to X)$ и $df(x) = dx = h$.

г. Всякая дифференцируемая при $x = c$ функция $f(x)$ непрерывна при $x = c$.

Действительно, в равенстве

$$f(c+h) - f(c) = f'(c)\,h + o(h)$$

для заданного $\varepsilon > 0$ возьмем $\delta > 0$ столь малым, чтобы при $|h| < \delta$ иметь $|o(h)| < \frac{\varepsilon}{2}$ и $\|f'(c)\| \cdot |h| < \varepsilon/2$;

тогда получится, что

$$|f(c+h)-f(c)| \leq \|f'(c)\|\cdot|h|+|o(h)|<\varepsilon,$$

откуда и следует непрерывность функции $f(x)$ при $x=c$.

I.32. а. Если функции $f(x):V\to Y$ и $g(x):V\to Y$ дифференцируемы при $x=c\in V$, то и $S(x)=f(x)+g(x)(V\to Y)$ дифференцируема при $x=c$; при этом

$$S'(c)=f'(c)+g'(c)$$

или, что то же,

$$dS(c)=df(c)+dg(c)$$

Действительно, из равенств

$$f(c+h)-f(c)=f'(c)h+o(h)$$

$$g(c+h)-g(c)=g'(c)h+o(h)$$

следует, что

$$S(c+h)-S(c)=f(c+h)-f(c)+g(c+h)-g(c)=$$

$$=[f'(c)+g'(c)]h+o(h)$$

откуда и вытекает требуемое.

б. Если функция $y(x):V\to Y$ дифференцируема при $x=a$, а A – линейный непрерывный оператор, действующий из пространства Y в (нормированное) пространство Z, то функция $z(x)=Ay(x)$ дифференцируема при $x=a$ и

$$z'(a)=Ay'(a),\ dz(a)=A\,dy(a)$$

Действительно, в данном случае

$$z(a+h)-z(a)=A[y(a+h)-y(a)]=$$

$$=A[y'(a)h+o(h)]=[Ay'(a)]h+o(h),$$

откуда и вытекает требуемое.

в. Пусть Y есть прямая сумма подпространств $Y_1,\dots,Y_n$ так что для любой функции $y(x):X\to Y$ имеются составляющие

$$y_1(x)=P_1y(x)(X\to Y_1),\dots,y_n(x)=P_ny(x)(X\to Y_n)$$

Утверждается, что дифференцируемость функции $y(x)$ при $x=a$ влечет дифференцируемость при $x=a$ и каждой из функций $y_k(x)$ $(k=1,\dots,n)$; это следует из непрерывности оператора P_k 1.17в) и б.

Обратное утверждение – если функции $y_1(x)(X\to Y_1),\dots,y_n(x)(X\to Y_n)$ дифференцируемы при $x=a$, то и функция

$y(x)=\{y_1(x),\dots,y_n(x)\}: X \to Y$ дифференцируема при $x=a$ — следует из a, поскольку

$$y(x)=y_1(x)+\dots+y_n(x) \qquad (1)$$

Производная $y'(a)$ есть непрерывный линейный оператор $X \to Y$, т.е. $y'(a) \in L(X,Y)$; аналогично $y'_k(a) \in L(X,Y_k)$ $(k=1,\dots,n)$. Пространство $L(X,Y)$ есть прямая сумма подпространств $L(X,Y_k)$ (1.14 ж), и из (1) немедленно следует, что составляющими элемента $y'(a) \in L(X,Y)$ являются величины $y'_k(a)$:

$$y'(a)=\{y'_1(a),\dots,y'_n(a)\} \qquad (2)$$

1.33. Производная и дифференциал от сложной функции.

а. Теорема. Пусть функция $y=y(x)$ действует из области G нормированного пространства X в нормированное пространство Y, $a \in G$, $y(a)=b$ и пусть функция $z=z(y)$ определена в окрестности точки b пространства Y и действует в нормированное пространство Z. Если функция $y(x)$ дифференцируема при $x=a$, а функция $z(y)$ дифференцируема при $y=b$, то сложная функция $\zeta(x)=z[y(x)]$, определенная в некоторой окрестности точки $a \in G$ и действующая в пространство Z, также дифференцируема при $x=a$ и

$$\zeta'(a)=z'(b)\cdot y'(a) \qquad (1)$$

Заметим, что $y'(a)$ есть линейный оператор, действующий из X в Y, а $z'(b)$ — линейный оператор, действующий из Y в Z, так что правая часть в формуле (1) имеет смысл, как линейный оператор, действующий из X в Z.

Для доказательства рассуждаем так:

$$\begin{aligned}\zeta(a+h)-\zeta(a) &= z[y(a+h)]-z[y(a)]=\\ &= z'[y(a)][y(a+h)-y(a)]+o[y(a+h)-y(a)]=\\ &= z'(b)[y'(a)h+o(h)]+o[y'(a)h+o(h)]=\\ &= z'(b)y'(a)h+o(h)\end{aligned} \qquad (2)$$

Значит, выражение $z'(b)\,y'(a)\,h$ составляет главную линейную часть в приращении функции $\zeta(x)$ при переходе от $x=a$ к $x=a+h$, что и требовалось проверить.

Заменяя a на x, b на $y(x)$ и переходя к операторам, можно записать полученную формулу в виде

$$\{z[y(x)]\}' = z'(y)\cdot y'(x) \qquad (3)$$

6. В частности, пусть пространства X, Y, Z конечномерны и их размерности равны соответственно n, m и p. Фиксируя каким-либо образом базисы в этих пространствах, функции $y(x)$ и $z(y)$ можно записать системами числовых равенств

$$\left.\begin{array}{ll} y_1 = y_1(x_1,\dots,x_n), & z_1 = z_1(y_1,\dots,y_m) \\ \dots\dots\dots & \dots\dots\dots \\ y_m = y_m(x_1,\dots,x_n), & z_p = z_p(y_1,\dots,y_m) \end{array}\right\} \qquad (4)$$

Операторам $y'(x)$ и $z'(y)$ отвечают матрицы Якоби (I.25 а)

$$y'(x) \cong \begin{Vmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial y_m}{\partial x_1} & \dots & \frac{\partial y_m}{\partial x_n} \end{Vmatrix}, \quad z'(y) \cong \begin{Vmatrix} \frac{\partial z_1}{\partial y_1} & \dots & \frac{\partial z_1}{\partial y_m} \\ \dots & \dots & \dots \\ \frac{\partial z_p}{\partial y_1} & \dots & \frac{\partial z_p}{\partial y_m} \end{Vmatrix} \qquad (5)$$

Функция $\zeta(x) = z[y(x)]$ по a имеет производную; соответствующая матрица Якоби в силу формулы (I) совпадает с произведением матриц (5):

$$\zeta'(x) \cong \begin{Vmatrix} \frac{\partial \zeta_1}{\partial x_1} & \dots & \frac{\partial \zeta_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial \zeta_p}{\partial x_1} & \dots & \frac{\partial \zeta_p}{\partial x_n} \end{Vmatrix} = \begin{Vmatrix} \frac{\partial z_1}{\partial y_1} & \dots & \frac{\partial z_1}{\partial y_m} \\ \dots & \dots & \dots \\ \frac{\partial z_p}{\partial y_1} & \dots & \frac{\partial z_p}{\partial y_m} \end{Vmatrix} \cdot \begin{Vmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial y_m}{\partial x_1} & \dots & \frac{\partial y_m}{\partial x_n} \end{Vmatrix}. \qquad (6)$$

Равенство (6) называется правилом умножения матриц Якоби. Его можно короче записать так:

$$\left\| \frac{\partial(\zeta_1,\dots,\zeta_p)}{\partial(x_1,\dots,x_n)} \right\| = \left\| \frac{\partial(z_1,\dots,z_p)}{\partial(y_1,\dots,y_m)} \right\| \cdot \left\| \frac{\partial(y_1,\dots,y_m)}{\partial(x_1,\dots,x_n)} \right\| \qquad (7)$$

В частности, для любых фиксированных $j=1,\dots,p$; $i=1,\dots,n$ мы имеем

$$\frac{\partial \zeta_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_i} \qquad (8)$$

Рассмотрим случай $n=1$, так что можно обозначить $x_1 = \dots = x_n = x$. Функции $y(x)$ и $\zeta(x)$ являются здесь функциями одного переменного x. Формула (8) приобретает вид

$$\frac{d\zeta_j}{dx} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \cdot \frac{dy_k}{dx} \qquad (9)$$

Мы видим, что и для дифференцирования функции одного переменного может потребоваться использование частных производных.

в. Заменяя в формуле (2) h на dx и вспоминая, что дифференциал функции есть главная линейная часть ее приращения, находим, что

$$d\zeta = z'(b) \cdot y'(a) \cdot dx$$

Но так как, с другой стороны, при данном dx и

$$dy = y'(a)\, dx,$$

то мы получаем

$$d\zeta = z'(b)\, dx$$

т.е. дифференциал функции от y не зависит от того, независимым ли переменным является ее аргумент y или же функцией от другого аргумента x. Это свойство называется инвариантностью диффференциала относительно замены переменного. Следует только иметь в виду, что в первом случае под dy понимается произвольное приращение аргумента y, тогда как во втором случае это значение дифференциала функции $y(x)$ для вектора dx.

I.34. Дифференциал обобщенного произведения.

а. Пусть имеются нормированные пространства X и Y и

определено произведение элементов $x \in X$ и $y \in Y$, т.е. билинейное непрерывное отображение $z = \langle x, y \rangle$ пространства $W = X + Y$ в некоторое пространство Z (I.18 в).

В силу непрерывности билинейной формы $\langle x, y \rangle$ существует постоянная $c > 0$ такая, что для любых x, y

$$|\langle x, y \rangle| \leq c\,|x| \cdot |y|$$

Утверждается, что функция $z = \langle x, y \rangle : W \to Z$ дифференцируема в каждой точке пространства W и

$$dz = \langle dx, y \rangle + \langle x, dy \rangle \qquad (I)$$

Действительно, придавая аргументу $\{x, y\} \in W$ приращение $\{dx, dy\}$ и используя билинейное свойство функции $\{x, y\}$, находим

$$\langle x + dx, y + dy \rangle - \langle x, y \rangle = \langle x, y \rangle + \langle dx, y \rangle + \langle x, dy \rangle +$$

$$+ \langle dx, dy \rangle - \langle x, y \rangle = \langle dx, y \rangle + \langle x, dy \rangle + \langle dx, dy \rangle$$

При этом первые два слагаемых справа линейны относительно смещения $\{dx, dy\}$, а выражение $\{dx, dy\}$ допускает оценку

$$|\langle dx, dy \rangle| \leq c\,|dx| \cdot |dy| \leq \frac{c}{2}\left(|dx|^2 + |dy|^2\right) =$$

$$= o\left(|dx| + |dy|\right)$$

Таким образом, приращение функции $\langle x, y \rangle$ допускает выделение главной линейной части $\langle dx, y \rangle + \langle x, dy \rangle$, что и доказывает наше утверждение.

Примерами применения этого правила, вместе с правилом дифференцирования сложной функции, являются формулы дифференцирования различных произведений, которые приводятся дальше.

6. Теорема. Пусть в области G пространства T заданы дифференцируемые функции $x = x(t), x : G \to X$ и $y = y(t), y : G \to Y$. Образуем, как в а , обобщенное произведение

$$\langle x(t), y(t) \rangle : G \to Z$$

которое называется обобщенным произведением функции $x(t)$ на функцию $y(t)$.

Утверждается, что функция $\zeta(t)=\langle x(t), y(t)\rangle$ дифференцируема в G и

$$d\zeta=\langle x'(t)dt, y(t)\rangle+\langle x(t), y'(t)dt\rangle \quad (2)$$

Действительно, функция $\zeta(t)$ может быть представлена как сложная функция

$$\{x,y\}=\{x(t),y(t)\}:G\to W$$

$$\zeta(t)=\langle x,y\rangle=z(\{x,y\}):W\to Z$$

Используя теорему о дифференциале сложной функции и результат а, получаем

$$d\zeta=\langle x'(t)dt, y(t)\rangle+\langle x(t), y'(t)dt\rangle$$

что и требуется.

Из (2) следует, что производную функции $\zeta(t)$ можно записать по формуле

$$\zeta'(t)=\langle x'(t), y(t)\rangle+\langle x(t), y'(t)\rangle \quad (3)$$

где слагаемые в правой части надо понимать как линейные операторы в пространстве T, действующие из T в Z по формулам

$$\langle x'(t), y(t)\rangle dt=\langle x'(t)dt, y(t)\rangle$$

$$\langle x(t), y'(t)\rangle dt=\langle x(t), y'(t)dt\rangle$$

в. Следствие. Если функция $x(t): G\subset T\to X$ дифференцируема при $t=c$ и $\lambda(t): G\subset T\to L(X,Y)$ — операторная функция, определенная в G и дифференцируемая при $t=c$, то произведение $g(t)=\lambda(t)\cdot x(t): G\subset T\to Y$ дифференцируемо при $t=c$ и

$$dg(c)=g'(c)\,dt=\lambda'(c)dt\cdot x(c)+\lambda(c)\cdot x'(c)dt \quad (4)$$

Оба слагаемых в правой части, как нетрудно видеть, принадлежат пространству Y.

г. Следствие. Если $\lambda(t): G\subset T\to R_1$ и $x(t): G\subset T\to R_1$ — две числовые функции, дифференцируемые при $t=c\in G$, то произведение $g(t)=\lambda(t)\cdot x(t)$ также дифференцируемо при $t=c$ и

$$g'(e) = x(e)\cdot a'(e) + a(e)\cdot x'(e) \qquad (5)$$

Правая и левая части равенства здесь являются линейными функционалами на пространстве T.

1.85. Функция x^{-1} и ее производная.

а. Предположим, что U и V – полные нормированные линейные пространства и все рассматриваемые отображения – линейные ограниченные операторы. Некоторые из линейных операторов $x: U \to V$ обратимы; их совокупность мы обозначим через G. На множестве G определена функция $x^{-1}: G \to L(V, U)$.

Теорема. В метрическом пространстве $L(U, V)$ (с обычной операторной метрикой) множество G есть область (открытое множество) и функция x^{-1} непрерывна на G.

Доказательство. Сначала рассмотрим частный случай этого утверждения для $U = V$. В этом случае операторы $x \in L(U)$ образуют нормированную алгебру (0.12.7). Рассмотрим единичный оператор e_U, очевидно, принадлежащий области G. Пусть $x = e - h$, где $|h| < 1$; покажем, что оператор x также принадлежит G. Положим

$$y = e + h + h^2 + \cdots + h^n + \cdots \qquad (I)$$

Так как $|h^n| \leq |h|^n$, то ряд (I) мажорируется числовой геометрической прогрессией и потому сходится в $L(U)$. Поскольку для абсолютно сходящихся рядов элементов нормированных пространств справедливы обычные правила перестановок, мы имеем

$$y\cdot x = (e + h + h^2 + \dots)(e - h) = (e + h + h^2 + \dots) - (h + h^2 + \dots) = e$$

$$x\cdot y = (e - h)(e + h + h^2 + \dots) = (e + h + h^2 + \dots) - (h + h^2 + \dots) = e$$

Таким образом, y есть левый и правый обратный для x, так что $x \in G$, что и утверждалось. Заметим здесь же, что из $x \to e$, т.е. $h \to 0$ из (I) следует, что и $y \to e$, так что операция взятия обратного непрерывна при $x = e$.

Теперь рассмотрим общий случай. Покажем, что у любого

$x \in G \subset L(U,V)$ имеется окрестность, принадлежащая множеству G. В поисках обратного элемента к $x+h$ при малом h мы умножим $x+h$ сначала на x^{-1}, убедимся, что результат – лежащий, очевидно, в $L(V)$ – при достаточно малых h будет близок к единице и, следовательно, будет обратимым в $L(V)$; а тогда, как мы увидим, будет обратимым в $L(U,V)$ и $x+h$. Проведем эту идею подробнее: для любого $h \in L(U,V)$ мы имеем $(x+h)x^{-1} = x \cdot x^{-1} + hx^{-1} = e_V + hx^{-1}$; поэтому при $|h| < 1/|x^{-1}|$ справедливо неравенство $|(x+h)x^{-1} - e_V| = |hx^{-1}| \leq |h||x^{-1}| < 1$, так что элемент $(x+h)x^{-1} \in L(V)$ отстоит от e_V менее чем на I. Но тогда оператор $(x+h)x^{-1}$ обратим в пространстве $L(V)$, так что существует оператор $z_h : V \to V$, для которого $(x+h)x^{-1}z_h = e_V$. Это означает, что $x^{-1}z_h \in L(V,U)$ есть правый обратный для $x+h$. Таким же образом, заменяя $(x+h)x^{-1}$ на $x^{-1}(x+h)$, мы можем доказать существование при $|h| < 1/|x^{-1}|$ и левого обратного для $x+h$. Теперь из I.I3в вытекает, что элемент $x+h$ при $|h| < 1/|x^{-1}|$ обратим. Мы видим, что множество G вместе со всяким своим элементом x содержит и все другие элементы совокупности $L(U,V)$, находящиеся от x на расстоянии меньшем $1/|x^{-1}|$; стало быть, G есть открытое множество. Так как $(x+h)x^{-1} \to e_V$ при $h \to 0$, то и $z_h = [(x+h)x^{-1}]^{-1} \to e_V$ при $h \to 0$, отсюда $(x+h)^{-1} = (x+h)^{-1} e_V = (x+h)^{-1}(x+h)x^{-1}z_h = x^{-1}z_h \to x^{-1}$ при $h \to 0$, и, следовательно, функция x^{-1} непрерывна на всей области своего определения.

б. Покажем, что в условиях а функция x^{-1} дифференцируема, и найдем ее дифференциал.

Из тождества

$$(x+h)[(x+h)^{-1} - x^{-1}]x = x - (x+h) = -h$$

имеем

$$(x+h)^{-1} - x^{-1} = -(x+h)^{-1} h x^{-1} = -x^{-1}hx^{-1} + o(h)$$

в силу непрерывности функции x^{-1} в области G (а).
Отсюда следует дифференцируемость функции x^{-1} на всей области ее определения и равенство

$$d(x^{-1}) = (x^{-1})'h = -x^{-1}hx^{-1}$$

Заметим, что вообще говоря, переставлять множители в полученном результате нельзя. Если даже $U = V$, так что формальная перестановка множителей будет иметь смысл, все равно ее делать нельзя из-за возможной некоммутативности операторов в $L(U)$. Лишь в случае $U = V = R$ множители можно переставлять безоговорочно. Функция x^{-1} здесь для каждого $x \neq 0$ есть оператор умножения на число $\frac{1}{x}$, который можно отождествить с функцией $\frac{1}{x}$; мы получаем здесь классическую формулу

$$d\left(\frac{1}{x}\right) = -\frac{dx}{x^2}$$

или

$$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}.$$

I.36. Дифференцируемость частного. Пусть $f(t)$ и $g(t)$ функции аргумента $t \in T$; значения функции $g(t)$ лежат в нормированном пространстве Y, значения функции $f(t)$ – в области $G \subset X$, где $X = L(U,V)$ есть нормированное пространство всех непрерывных линейных операторов, действующих из нормированного пространства U в нормированное пространство V, а $G \subset X$ состоит из обратимых операторов, так что для каждого $x \in G$ определен оператор $f^{-1}(t) \in L(V,U)$. Пусть далее задано обобщенное произведение $X \times Y \to Z$ (I.18.в). Тогда частным функций $f(t)$ и $g(t)$ называется обобщенное произведение вида $\langle f^{-1}(t), g(t)\rangle$.

Определенное таким образом частное есть композиция более простых функций $f(t)$, $g(t)$, x^{-1}, $\langle\ \rangle$; по теореме о дифференцируемости сложной функции онон представляет собой дифференцируемую функцию от t, если дифференцируемы $f(t)$ и $g(t)$. Формулу для дифференциала частного нетрудно написать, используя I.33а, I.33в, I.34а и I.35б:

$$d\langle f^{-1}(t), g(t)\rangle = \langle d f^{-1}(t), g(t)\rangle + \langle f^{-1}(t), dg(t)\rangle =$$

$$= \langle -f^{-1}(t)\, df(t)\, f^{-1}(t), g(t)\rangle + \langle f^{-1}(t), g'(t)\,dt\rangle =$$

$$= \langle -f^{-1}(t)\, f'(t)\,dt\, f^{-1}(t), g(t)\rangle + \langle f^{-1}(t), g'(t)\,dt\rangle$$

Если $f(t)$ и $g(t)$ числовые функции и $\langle\ \ \rangle$ означает обычное (числовое) произведение, мы получаем формулу классического анализа

$$d\,\frac{g(t)}{f(t)} = -\frac{f'(t)\,g(t)\,dt}{f^2(t)} + \frac{g'(t)\,dt}{f(t)} = \frac{g'(t)\,f(t) - f'(t)\,g(t)}{f^2(t)}\,dt$$

Заметим, что эта формула все еще остается более общей, чем в классическом анализе, поскольку пространство T произвольно.

§ I.4. Теорема о конечном приращении.

I.4I. а. В классическом анализе теорема о конечном приращении (теорема Лагранжа) имеет следующую формулировку:

Если $f(x)$ числовая функция, заданная и дифференцируемая на отрезке $a \le x \le b$, то существует точка $c \in [a, b]$ такая, что

$$f(b) - f(a) = f'(c)\,(b - a) \qquad (I)$$

Следствием соотношения (I) является часто используемая оценка

$$|f(b) - f(a)| \le \sup_{a \le x \le b} |f'(x)| \cdot (b - a) \qquad (2)$$

б. Мы желаем найти аналоги (I) и (2) для дифференцируемой функции $y = f(x) : G \subset X \to Y$, где X и Y нормированные пространства. Заметим, что приведенная формулировка теоремы Лагранжа не переносится буквально даже на случай $X = R_1$, $Y = R_2$. Действительно, для функции

$$y = f(x) = \cos x \cdot e_1 + \sin x \cdot e_2, \quad 0 \le x \le 2\pi,$$

где e_1 и e_2 базисные векторы в R_2, мы имеем

$f(0) - f(2\pi) = 0$, $f'(x) = -\sin x \cdot e_1 + \cos x \cdot e_2$

и при $\alpha=0$, $\ell=2\pi$ равенство (I) заведомо не выполняется ни при каком $c\in[0,2\pi]$, поскольку $f'(x)$ не обращается в 0. Таким образом, если (I) и возможно перенести на общий случай, то с существенным изменением в формулировке; мы ниже укажем это изменение.

в. Что касается оценки (2), то ее естественное обобщение оказывается справедливым. Вначале мы рассмотрим случай функции одного вещественного переменного:

Лемма. Если функция $y=\varphi(t):[\alpha,\beta]\to Y$ дифференцируема на $[\alpha,\beta]$, то оценка (2), в которой положено $a=\alpha$, $\ell=\beta$, $f=\varphi$ справедлива.

Доказательство. Выберем произвольно $\varepsilon>0$ и рассмотрим множество Q всех $t\in[\alpha,\beta]$, для которых справедливо неравенство

$$|\varphi(t)-\varphi(\alpha)|\leq(M+\varepsilon)(t-\alpha) \qquad (3)$$

где $M=\sup\limits_{\alpha\leq t\leq\beta}|\varphi'(t)|$. Так как функция $\varphi(t)$ непрерывна, то Q – замкнутое множество в $[\alpha,\beta]$. Оно содержит некоторую окрестность точки α, поскольку в силу дифференцируемости функции $\varphi(t)$ при $t=\alpha$ существует такое δ, что при $\alpha\leq t\leq t+\delta$

$$|\varphi(t)-\varphi(\alpha)|\leq|\varphi'(\alpha)(t-\alpha)+\varepsilon(t-\alpha)|\leq(M+\varepsilon)(t-\alpha)$$

Мы покажем, что Q совпадает со всем отрезком $[\alpha,\beta]$. Если это не так, то пусть $Z=[\alpha,\beta]-Q$ и $\xi=\inf\limits_{z\in Z} z$. Поскольку Z открыто, мы имеем $\xi<\beta$ и по доказанному $\xi>\alpha$. В силу дифференцируемости функции $\varphi(t)$ при $t=\xi$ существует такое δ, что при $\xi-\delta<t<\xi$ выполняется неравенство

$$|\varphi(t)-\varphi(\xi)|\leq|\varphi'(t)||t-\xi|+\frac{\varepsilon}{2}|t-\xi|\leq\left(M+\frac{\varepsilon}{2}\right)(\xi-t) \qquad (4)$$

С другой стороны, поскольку $t\in Q$, выполняется неравенство (3). Из (3) и (4) имеем

$$|\varphi(\xi)-\varphi(\alpha)|\leq|\varphi(\xi)-\varphi(t)|+|\varphi(t)-\varphi(\alpha)|\leq$$

$$\leq M(\xi-\alpha)+\varepsilon(t-\alpha)+\frac{\varepsilon}{2}(\xi-t)<(M+\varepsilon)(\xi-\alpha)$$

Так как получившееся неравенство - строгое, то оно выполняется и для всех t в некоторой окрестности точки ξ. А это противоречит тому, что в любой окрестности точки ξ по построению имеются точки из множества Z. Таким образом, неравенство (3) имеет место при всех $t\in[\alpha,\beta]$. Полагая $t=\beta$, получаем

$$|\varphi(\beta)-\varphi(\alpha)|\leq(M+\varepsilon)(\beta-\alpha)$$

Так как здесь $\varepsilon>0$ произвольно мало, то имеет место и неравенство (2), что и требовалось.

г. Если предположить в условиях в, что функция $\varphi'(t)$ не только существует на $[\alpha,\beta]$, но и непрерывна, или хотя бы кусочно непрерывна, то для разности $\varphi(\beta)-\varphi(\alpha)$ можно воспользоваться представлением (см. 012.53 в)

$$\varphi(\beta)-\varphi(\alpha)=\int_{\alpha}^{\beta}\varphi'(t)\,dt$$

откуда сразу вытекает и оценка (2). Это же представление позволяет дать правильную формулировку и обобщению формулы (I). Именно, можно написать (см. подробнее в 012.53), что

$$\varphi(\beta)-\varphi(\alpha)=Q(\beta-\alpha),$$

где Q есть некоторая точка из замкнутой выпуклой оболочки значений $\varphi'(t)$ на $[\alpha,\beta]$.

д. Мы можем перейти теперь к самому общему случаю.

Теорема о конечном приращении. Пусть дана дифференцируемая функция $y=f(x):G\subset X\to Y$. В области G возьмем две точки a и b так, что соединяющий их отрезок целиком лежит в G. Тогда имеет место неравенство

$$|f(b)-f(a)|\leq\sup_{x\in G}|f'(x)(b-a)| \qquad (5)$$

Доказательство. Отрезок, соединяющий точки a и b, есть годограф функции $x(t)=(1-t)a+tb:[0,1]\to G$. Эта функция дифференцируема и $x'(t)=b-a$. Введем

функцию

$$\varphi(t) = f[x(t)] : [0,1] \to Y$$

Имеем $\varphi(0) = f(a)$, $\varphi(1) = f(b)$.

Как композиция дифференцируемых функций, функция $\varphi(t)$ также дифференцируема и

$$\varphi'(t) = f'(x)\cdot x'(t) = f'(x)\cdot(b-a)$$

Для функции $\varphi(t)$ по лемме в справедлива оценка (2) (где положено $a=0, b=1$):

$$|\varphi(1) - \varphi(0)| \leq \sup_{0\leq t\leq 1} |\varphi'(t)|$$

Заменяя здесь $\varphi(1)$ на $f(b)$, $\varphi(0)$ на $f(a)$, $\varphi'(t)$ на $f'(x)(b-a)$, и переходя от множества значений $f'(x)(b-a)$ на отрезке, соединяющим точки a и b в области G, к множеству значений этой функции во всей области G, получаем неравенство (5), что и требуется.

Неравенство (5) используется и в более грубой, но более простой форме

$$|f(b) - f(a)| \leq \sup_{x\in G} \|f'(x)\|\cdot|b-a| \qquad (6)$$

Неравенство (5) дает более сильный результат, чем неравенство (6), например, в случае, когда $X = R_n$, $Y = R_1$ и вектор $\operatorname{grad} y(x)$ направлен под углом к вектору $b-a$, близким к прямому.

е. Уточненная теорема о конечном приращении. В предположениях д справедливо также неравенство

$$|f(b) - f(a) - f'(a)(b-a)| \leq \sup_{x\in G} |(f'(x) - f'(a))(b-a)| \qquad (7)$$

Доказательство. Рассмотрим функцию

$$g(x) = f(x) - f'(a)(x-a)$$

Она дифференцируема вместе с функцией $f(x)$, и по I.32а-б

$$g'(x) = f'(x) - f'(a)$$

Применяя для функции $g(x)$ неравенство (5), находим

$$|f(b)-f(a)-f'(a)(b-a)| \equiv |g(b)-g(a)| \leq$$

$$\leq \sup_{x\in G} |g'(x)(b-a)| = \sup_{x\in G} |(f'(x)-f'(a))(b-a)|$$

что и требовалось.

Неравенство (7) также используется и в более грубой форме

$$|f(b)-f(a)-f'(a)(b-a)| \leq \sup_{x\in G} \|f'(x)-f'(a)\| \cdot |b-a| \quad (8)$$

I.42. В дальнейшем будут даны простейшие приложения теоремы о конечном приращении.

а. Теорема. Если в шаре $V=\{x\in X: |x-a|\leq r\}$ мы имеем $y'(x)\equiv 0$, то $y(x)$ есть постоянная.

Действительно, для любого $b\in V$ и отрезка L_b, соединяющего точку a с точкой b, имеет место неравенство I.41 (6):

$$|y(b)-y(a)| \leq \sup_{x\in L_b} \|y'(x)\| \cdot |b-a| = 0,$$

откуда $y(b)=y(a)$.

б. Следствие. Если в шаре $V=\{x\in X: |x-a|\leq r\}$ функции $y_1(x)$ и $y_2(x)$ имеют совпадающие производные, то $y_1(x)$ и $y_2(x)$ отличаются в этом шаре на аддитивную постоянную.

Действительно, $y_1(x)-y_2(x)$ имеет производную, равную 0, и можно применить а .

в. Следствие. Если в шаре $V=\{x\in X: |x-a|\leq r\}$ оператор $y'(x)$ постоянен, $y'(x)\equiv y'(a)$, то

$$y(x) = y'(a)(x-a) + y(a)$$

Для доказательства следует действовать тем же путем, что и в а, но вместо неравенства I.41 (6) использовать неравенство I.41 (7). Можно обойтись и без неравенства I.41 (7): функция $g(x)=y'(a)(x-a)+y(a)$ имеет производную $g'(x)=y'(a)$; поэтому, согласно б, функции $g(x)$ и $y(x)$ отличаются на постоянную; полагая $x=a$, находим, что эта постоянная равна 0.

I.43. Производная и условие Липшица.

а. Функция $y = y(x)$ $(G \subset X \to Y)$ по определению удовлетворяет в шаре $V = \{x \in X : |x-a| \leq r\} \in G$ условию Липшица, если существует такая постоянная $C > 0$, что при всех $x_1 \in V$, $x_2 \in V$ выполняется неравенство

$$|y(x_1) - y(x_2)| \leq C|x_1 - x_2| \qquad \text{(I)}$$

Предположим, что функция $y(x)$ дифференцируема в шаре V, и пусть

$$\sup_{x \in V} \|y'(x)\| = B$$

Тогда в силу I.4I (б) мы имеем

$$|y(x_1) - y(x_2)| \leq B|x_1 - x_2|$$

т.е. функция $y(x)$ в шаре V удовлетворяет условию Липшица (I) с постоянной B.

Обратно, если функция $y(x)$ имеет непрерывную производную $y'(x)$ и удовлетворяет в шаре V условию Липшица (I), то можно утверждать, что $\|y'(x)\| \leq C$. Действительно, пусть $|x-a| = \rho < r$; для заданного $\varepsilon > 0$ найдем $\delta > 0$, $\delta < r - \rho$ так, чтобы для любого x_1 с $|x - x_1| \leq \delta$ имело место неравенство

$$|y(x_1) - y(x) - y'(x)(x_1 - x)| < \varepsilon|x_1 - x|$$

Так как по предположению $|y(x_1) - y(x)| \leq C|x_1 - x|$ то мы имеем при $|x_1 - x| < \delta$

$$|y'(x)(x_1 - x)| \leq (C + \varepsilon)|x_1 - x|$$

откуда для нормы оператора $y'(x)$ получается неравенство

$$\|y'(x)\| \leq C + \varepsilon$$

поскольку $\varepsilon > 0$ произвольно, это означает, что $\|y'(x)\| \leq C$ и утверждение доказано.

Отметим, что из выполнения одного только условия Липшица (I) дифференцируемость функции $y(x)$, вообще говоря, не следует (даже в случае $X = R_1$; пример: $y = |x|$ $(R_1 \to R_1)$).

б. Интересный частный случай получается при условии $\sup\limits_{x \in V} \|y'(x)\| < 1$. По доказанному это означает, что

функция $y(x)$ удовлетворяет условию Липшица с постоянной C, меньшей 1. Другими словами, отображение $y=y(x)$ уменьшает расстояния: расстояние между точками $y(x)$ и $y(\alpha)$ строго меньше расстояния между точками x и α. Если к тому же значения функции $y(x)$ принадлежат области V, а $\sup\|y'(x)\|\leq\Theta<1$, то отображение $y(x): V\to V$ оказывается сжимающим

$$|y(x')-y(x'')|\leq\Theta|x'-x''| \qquad (2)$$

в. Вообще сжимающие отображения определяются даже в метрических пространствах; именно, если M метрическое пространство, то отображение $y=y(x): M\to M$ называется сжимающим, если существует такая постоянная Θ, $0<\Theta<1$, что

$$\rho(y(x_1), y(x_2))\leq\Theta\,\rho(x_1,x_2)$$

для любых двух точек x_1 и x_2 из M.

В полном метрическом пространстве M сжимающее отображение $y(x)$ всегда обладает неподвижной точкой – такой точкой z, что $y(z)=z$. Её можно построить, как предел последовательности $x_1,\dots,x_n,\dots$ точек пространства M, где x_1 берется в M произвольно, $x_2=y(x_1)$, $x_3=y(x_2)$, $\dots$, $x_{n+1}=y(x_n)$, $\dots$. Действительно, мы имеем полагая

$$\rho(x_1,x_2)=c,$$
$$\rho(x_2,x_3)=\rho(y(x_1),y(x_2))\leq\Theta\,\rho(x_1x_2)=\Theta\cdot c$$
$$\rho(x_3,x_4)=\rho(y(x_2),y(x_3))\leq\Theta\,\rho(x_2x_3)\leq\Theta^2\cdot c$$
$$\dots\dots\dots$$
$$\rho(x_n,x_{n+1})\leq\Theta^{n-1}\cdot c,$$
$$\dots\dots\dots$$

Отсюда для любых m и n, $m<n$

$$\rho(x_m,x_n)\leq\rho(x_m,x_{m+1})+\dots+\rho(x_{n-1},x_n)\leq$$
$$\leq(\Theta^{m-1}+\dots+\Theta^{n-2})\cdot c$$

Выражение в скобках есть отрезок геометрической прогрессии с знаменателем $\theta < 1$ и поэтому стремится к 0 при $m, n \to \infty$. Таким образом, последовательность $x_1, x_2, \ldots$ фундаментальна в M. Так как M полно, то эта последовательность сходится; пусть $z = \lim_{n\to\infty} x_n$.

Мы имеем далее

$$\rho(z, y(z)) \le \rho(z, x_n) + \rho(x_n, y(z)) = \rho(z, x_n) + \rho(y(x_{n-1}), y(z)) \le \rho(z, x_n) + \theta\rho(x_{n-1}, z) \to 0 \text{ при } n\to\infty,$$

откуда $\rho(z, y(z)) = 0$, $y(z) = z$, так что z действительно неподвижная точка отображения y.

Добавим к сказанному, что неподвижная точка у сжимающего отображения может быть лишь единственной: если бы z_1 и z_2 были бы неподвижными, то мы имели бы

$$\rho(z_1, z_2) = \rho(y(z_1), y(z_2)) \le \theta\rho(z_1, z_2)$$

что возможно лишь при $\rho(z_1, z_2) = 0$, т.е. $z_1 = z_2$.

г. Пусть, как и в б, в шаре $V \subset X$ имеется дифференцируемая функция $y(x)$: $\sup \|f'(x)\| = \theta < 1$, отображающая этот шар в себя. Неравенство (2) продолжается по непрерывности в замкнутый шар $\overline{V}$; таким образом, отображение $y(x)$ является сжимающим в полном метрическом пространстве $\overline{V}$. В силу в отображение $y(x)$ в шаре $\overline{V}$ обладает неподвижной точкой, притом единственной.

д. Можно указать простое достаточное условие того, чтобы функция $y(x)$ $(V \subset X \to X)$, удовлетворяющая в шаре $V = \{x \in X : |x-a| \le r\}$ условию Липшица с постоянной $\theta < 1$, переводила этот шар в себя. Именно, потребуем, чтобы выполнялось неравенство $|y(a) - a| \le (1-\theta)r$. Тогда для любого $x \in V$ мы будем иметь

$$|y(x) - a| \le |y(x) - y(a)| + |y(a) - a| \le \theta|x-a| + (1-\theta)r = r,$$

так что все значения функции $y(x)$ при $x \in V$ лежат в шаре V.

е. Комбинируя результаты в, г и д, получаем теорему:

Теорема. Если в шаре $V = \{x \in X : |x-a| \le r\}$ функция $y(x) : V \to X$ дифференцируема и при некотором

$\theta<1$ выполняются неравенства

$$\sup_{x\in V}\|\varphi'(x)\|\le\theta,\quad |\varphi(a)-a|\le(1-\theta)r,$$

то в шаре V существует и единственна точка x_0, для которой $\varphi(x_0)=x_0$.

Этим способом доказательства существования неподвижных точек мы в дальнейшем воспользуемся.

I.44. а. Многие теоретические и вычислительные задачи приводятся к решению уравнения вида

$$f(x)=0 \qquad (I)$$

где $f(x): G\subset X\to Y$ дифференцируемая функция. Уравнение (I) в расшифрованном виде может представлять собой систему алгебраических или трансцендентных уравнений с n переменными, дифференциальное или интегральное уравнение и пр. Существует метод решения этого уравнения (не всегда приводящий к цели), который называется методом Ньютона или методом касательных; в общей форме он был сформулирован Л.В.Канторовичем в 1948 г. Пусть a какое-то значение x; если $f(a)=0$, то a и есть искомый корень, так что будем предполагать, что $f(a)\neq 0$. Если бы функция $f(x)$ была линейной, то для любого x_1 мы имели бы

$$f(x_1)-f(a)=f'(a)(x_1-a)$$

и если бы x_1 было искомым корнем, то мы получили бы для x_1 уравнение

$$-f(a)=f'(a)(x_1-a) \qquad (2)$$

Предположим далее, что оператор $f'(a): X\to Y$ обратим. Тогда уравнение (2) можно было бы разрешить и мы бы получили

$$x_1=a-[f'(a)]^{-1}\cdot f(a) \qquad (3)$$

В общем случае, когда $f(x)$ не есть линейная функция, величина x_1 вообще говоря не будет корнем уравнения $f(x)=0$. Но, возможно, значение $f(x_1)$ ближе к нулю, чем $f(a)$.

Формула (3) подсказывает, что следует рассмотреть последовательность

$$x_{n+1}=x_n-[f'(x_n)]^{-1}f(x_n),\ x_0=a, \quad (4)$$

или даже более грубую, но более простую

$$x_{n+1}=x_n-[f'(a)]^{-1}f(x_n),\ x_0=a \quad (5)$$

Выясним, в каком случае последовательность (5) будет сходиться. Для этого зададимся каким-либо значением $r>0$ и рассмотрим в шаре $V=\{x\in X: |x-a|\leq r\}$ функцию

$$y(x)=x-[f'(a)]^{-1}\cdot f(x): V\to X \quad (6)$$

и спросим себя при каких условиях для нее будут выполнены условия теоремы I.43 е. Мы имеем

$$y'(x) = E-[f'(a)]^{-1}\cdot f'(x) =$$

$$= -[f'(a)]^{-1}\cdot[f'(x)-f'(a)],$$

$$y(a)-a = -[f'(a)]^{-1}f(a)$$

Теорема. При выполнении условий

$$\sup_{|x-a|\leq r}\|[f'(a)]^{-1}[f'(x)-f'(a)]\|=\theta<1 \quad (7)$$

$$|[f'(a)]^{-1}f(a)|\leq(1-\theta)\cdot r \quad (8)$$

у уравнения (I) имеется решение в шаре V и это решение получается, как предел последовательности (5).

Доказательство. Из условия (7) следует, что в шаре выполняется неравенство $\|y'(x)\|\leq\theta<1$, а из условия (8) – что $|y(a)-a|\leq(1-\theta)\cdot r$. Применяя I.43е, получаем, что у отображения (6) имеется неподвижная точка, т.е. такая точка $z\in M$, что

$$z=z-[f'(a)]^{-1}\cdot f(z)$$

и, следовательно, $f(z)=0$; таким образом точка z является искомым корнем уравнения (I). Неподвижная точка

сжимающего отображения получается, как предел последовательности $x_{n+1} = y(x_n)$, что и завершает доказательство.

<u>б</u>. Метод Ньютона иллюстрируется на рис. I.44-I, где сходимость последовательных приближений очевидна. Но, конечно, легко нарисовать картинки с функцией $f(x)$, для которой последовательные приближения по методу Ньютона не сходятся к пределу.

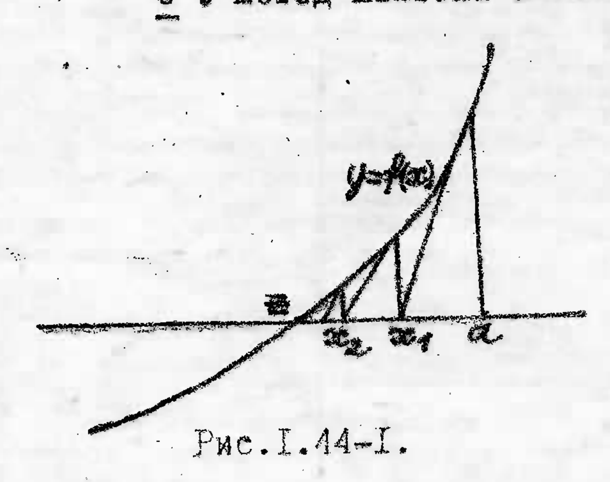


Рис.I.44-I.

На рис. I.44-2 представлена такая ситуация. Разумеется, в этом случае условия а не выполнены, и представляет интерес выяснение общего вопроса, в какой области пространства X для данной функции $f(x)$ они выполняются. (для выпуклой числовой функции $f(x)$ на числовой оси, как видно из рис. I.44-I, метод Ньютона сходится для любого начального значения $a > z$ и для некоторых $a < z$)

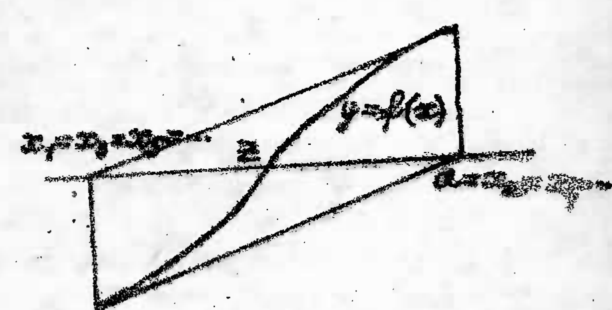


Рис.I.44-2. Уравнение $f(x)=0$ имеет корень z, но последовательность a, x_1, x_2, ... не имеет предела.

<u>в</u>. В общем случае мы можем лишь утверждать, что <u>область сходимости метода Ньютона содержит некоторую окрестность точки</u> z, если $f'(x)$ непрерывна при $x = z$. Действительно, выберем вначале такое $\rho > 0$, чтобы при $|x - z| \leq \rho$ удовлетворялось неравенство

$$\sup_{|x-z|\leq\rho} \| [f'(x)]^{-1} \| \cdot \sup_{|x-z|\leq\rho} \| f'(x) - f'(z) \| < \frac{1}{4} \qquad (9)$$

а затем такое $\delta \leq \frac{\rho}{2}$, чтобы при $|x - z| \leq \delta$ выполнялось неравенство

$$| [f'(x)]^{-1} \cdot f(x) | < \frac{\rho}{4} \qquad (10)$$

Первое неравенство обеспечивается предположенной непрерывностью функции $f'(x)$ при $x = z$, второе – непрерывностью $f(x)$ (и условием $f(z) = 0$). Теперь мы утверждаем, что для любого a с $|a - z| \leq \delta$ последовательность (5) сходится. Действительно, для такой точки a выполняется, во первых, неравенство (10), совпадающее с (8) при $\theta = \frac{1}{2}$ и $r = \frac{\rho}{2}$; а во вторых, поскольку шар $|a - x| \leq \frac{\rho}{2}$ содержится в шаре $|x - z| \leq \frac{\rho}{2}$, и неравенство

$$\| [f'(a)]^{-1} [f'(x) - f'(a)] \| \leq$$

$$\leq \| [f'(a)]^{-1} \cdot \{ [f'(x) - f'(z)] + [f'(z) - f'(a)] \} \| \leq$$

$$\leq \| [f'(a)]^{-1} \cdot [f'(x) - f'(z)] \| + \| [f'(a)]^{-1} \cdot [f'(z) - f'(a)] \| < \frac{1}{2}$$

Таким образом, в шаре $|x - a| \leq \frac{\rho}{2}$ выполняются оба неравенства (8) и (7), при $\theta = \frac{1}{2}$ и $r = \frac{\rho}{2}$. Следовательно, по теореме a, последовательность (5) сходится при выполнении наших условий.

Вопросам о сходимости метода Ньютона и применениям этого метода посвящена большая литература. См. например, Л.В.Канторович и Г.П.Акимов, Функциональный анализ в нормированных пространствах, Физматгиз 1959, гл. 18.

I.45. Теорема. Пусть в шаре $V = \{x \in X : |x - a| \leq r\}$ задана последовательность дифференцируемых функций $y_1(x)$, $y_2(x), \ldots$, со значениями в полном пространстве Y, производные которых $y_1'(x), y_2'(x), \ldots$ $(X \to L(X, Y))$ непрерывны и сходятся равномерно в V к некоторой функции $g(x)$ $(X \to L(X, Y))$. Если векторы $y_1(a), y_2(a), \ldots$, пространства Y имеют предел, то функции $y_1(x), y_2(x), \ldots$ сходятся равномерно в V к некоторой функции $y(x)$ (со значениями в Y), дифференцируемой внутри шара V, причем $y'(x) = g(x)$.

Доказательство. Напишем формулу I.41 (6), в которой y заменим на $y_n - y_m$ и b на x:

$$| [y_n(x) - y_m(x)] - [y_n(a) - y_m(a)] | \leq$$

$$\leq \sup_{x \in V} \| y_n'(x) - y_m'(x) \| \cdot |x - a|$$

В силу предположенной равномерной сходимости функций $y_n'(x)$ правая часть стремится к 0 при $n \to \infty$, $m \to \infty$. Используя сходимость векторов $y_n(a)$, мы получаем, что разность $y_n(x) - y_m(x)$ стремится к 0 при $n \to \infty$, $m \to \infty$ равномерно по $x \in V$. Таким образом, последовательность $y_n(x)$ фундаментальна в Y, и так как Y полно, то существует предельная функция $y(x) = \lim_n y_n(x)$. Эта функция $y(x)$ непрерывна вследствие равномерной сходимости $y_n(x)$ (I.16 б). Далее, напишем формулу I.41 (8) для функции $y_n(x)$ в шаре радиуса ρ с центром в точке b, $|b-a| \le r-\rho$:

$$|y_n(x) - y_n(b) - y_n'(b)(x-b)| \le \sup_{|\xi-b|\le\rho} \|y_n'(\xi) - y_n'(b)\| \cdot |x-b| \qquad (I)$$

Для заданного $\varepsilon > 0$ можно выбрать $\delta > 0$, $\delta \le \rho$ так, чтобы иметь для всех достаточно больших $n = N, N+1, \ldots$

$$\sup_{|\xi-b|\le\delta} \|y_n'(\xi) - y_n'(b)\| \le \varepsilon$$

Чтобы убедиться в этом, достаточно записать

$$y_n'(\xi) - y_n'(b) = [g(\xi) - g(b)] - [g(\xi) - y_n'(\xi)] + [g(b) - y_n'(b)]$$

и воспользоваться тем, что сходимость $y_n'(x)$ к $g(x)$ равномерна, а функция $g(x)$ непрерывна.

Заменяя в (I) ρ на δ и переходя к пределу при $n \to \infty$ находим, что при всех x, $|x-b| \le \delta$

$$|y(x) - y(b) - g(b)(x-b)| \le \varepsilon |x-b|$$

это и показывает, что $y(x)$ дифференцируема при $x=b$ и $y'(b)=g(b)$. Теорема доказана.

В теореме I.45 шар V можно заменить на связную область (т.е. такую область V, в которой любые две точки x_0 и x_1 можно соединить конечно-звенной ломаной).

I.46. Производные по подпространству.

а. Согласно определению, если функция $y=y(x)$ $(G\subset X\to Y)$ дифференцируема при $x=c\in G$, то линейный оператор $y'(c)$ определен на всем пространстве X и главная линейная часть приращения функции $y(x)$ при приращении независимого переменного, равном h, есть $y'(c)h$. Но можно рассматривать вопрос о дифференцируемости функции $y(x)$, ограничиваясь приращениями независимого переменного только в пределах некоторого линейного подпространства $X_1\subset X$.

Будем называть функцию $y(x)$ $(G\subset X\to Y)$ дифференцируемой по подпространству $X_1\subset X$, если приращение $y(x)$ при переходе из точки $x=c$ в точку $x=c+h, h\in X_1$ допускает выделение главной линейной части относительно $h\in X_1$:

$$y(c+h)-y(c)=\mathcal{D}_1(c)h+o(h),$$

где $\mathcal{D}_1(c)$ есть линейный оператор, определенный на подпространстве X_1. Этот оператор называется оператором частного дифференцирования по подпространству X_1. Иногда аргумент x в пределах подпространства X обозначают каким-либо специальным символом, например, x_1 (сохраняя обозначение x для векторов всего пространства X), тогда для оператора $\mathcal{D}_1(c)$ используют соответственно обозначение $\frac{\partial y(c)}{\partial x_1}$.

б. Выясним, например, что означает дифференцируемость функции $y(x)$ $(G\subset R_n\to Y)$ по одномерному подпространству X_k, определенному координатной осью x_k. В этом случае $h=(0,\ldots,h_k,\ldots,0)$ и справедливо равенство

$$y(c+h)-y(c)=y(c_1,\ldots,c_k+h_k,\ldots,c_n)-$$
$$-y(c_1,\ldots,c_k,\ldots,c_n)=\mathcal{D}_k(c)h_k+o(h_k),$$
$$\mathcal{D}_k(c): X_k\to Y.$$

Но это равносильно существованию обычной частной производной $\frac{\partial y(c)}{\partial x_k}$ (I.22 б), которая как раз и совпадает с величиной $\mathcal{D}_k(c)$.

в. Дифференцируемость функции $y(x): R_n \to R_m$ в точке $x=c$ по подпространству R_k, порожденному первыми k базисными векторами пространства R_n, приводит к существованию всех частных производных, входящих в матрицу.

$$\frac{\partial(y_1,\dots,y_m)(c)}{\partial(x_1,\dots,x_k)} = \begin{Vmatrix} \frac{\partial y_1(c)}{\partial x_1} & \dots & \frac{\partial y_1(c)}{\partial x_k} \\ \dots & \dots & \dots \\ \frac{\partial y_m(c)}{\partial x_1} & \dots & \frac{\partial y_m(c)}{\partial x_k} \end{Vmatrix}$$

Линейный оператор $(R_k \to R_m)$, соответствующий этой матрице, и есть частная производная функции $y(x)$ по подпространству R_k.

г. Очевидно, если функция $y(x)$ дифференцируема в точке c в исходном смысле I.23, она будет дифференцируемой в этой точке и по любому подпространству $X_1 \subset X$, причем соответствующий линейный оператор $\frac{\partial y(c)}{\partial x_1}$ "частного дифференцирования" есть просто сужение оператора $\frac{\partial y(c)}{\partial x}$ на подпространство X.

I.47. Из дифференцируемости функции $y(x)$ по (истинному) подпространству $X_1 \subset X$, вообще говоря, не следует ее дифференцируемость по всему X. А, например, для $X = R_n$ из дифференцируемости функции $y(x)$ в точке $x=c$ даже по любому подпространству размерности $k<n$ не вытекает ее дифференцируемость по всему пространству R_n (см. задачу 24). Справедлива следующая теорема:

а. Теорема. Если пространство X есть нормированная прямая сумма подпространств X_1 и X_2, а функция $y(x): G \subset X \to Y$ дифференцируема в некоторой окрестности точки $c \in G$ по подпространствам X_1 и X_2, причем производные $\frac{\partial y(x)}{\partial x_1}$ и $\frac{\partial y(x)}{\partial x_2}$ непрерывны в точке c, то функция $y(x)$ дифференцируема в точке c и по всему пространству X.

Доказательство. Для любого $h \in X$ можно написать $h = h_1 + h_2, h_1 \in X_1, h_2 \in X_2$, причем соотношение $h \to 0$ эквивалентно $h_1 \to 0$, $h_2 \to 0$. Далее, имеем при достаточно малых h

$$y(c+h) - y(c) = [y(c+h_1+h_2) - y(c+h_1)] + [y(c+h_1) - y(c)]$$

и по теореме о конечном приращении I.41.е мы находим

$$\left| y(c+h) - y(c) - \frac{\partial y(c+h_1)}{\partial x_2} h_2 - \frac{\partial y(c)}{\partial x_1} h_1 \right| \le$$

$$\le \sup_{0 \le t \le 1} \left\| \frac{\partial y(c+h_1+t h_2)}{\partial x_2} - \frac{\partial y(c+h_1)}{\partial x_2} \right\| \cdot |h_2| +$$

$$+ \sup_{0 \le \tau \le 1} \left\| \frac{\partial y(c+\tau h_1)}{\partial x_1} - \frac{\partial y(c)}{\partial x_1} \right\| \cdot |h_1|$$

В силу предположенной непрерывности производных $\frac{\partial y(x)}{\partial x_1}$ и $\frac{\partial y(x)}{\partial x_2}$ в точке c далее можно написать

$$\frac{\partial y(c+h_1)}{\partial x_2} = \frac{\partial y(c)}{\partial x_2} + o(1),$$

$$\frac{\partial y(c+h_1+t h_2)}{\partial x_2} - \frac{\partial y(c+h_1)}{\partial x_2} = o(1),$$

$$\frac{\partial y(c+\tau h_1)}{\partial x_1} - \frac{\partial y(c)}{\partial x_1} = o(1),$$

где $o(1)$ стремится к 0 при $h_1 \to 0$, $h_2 \to 0$. Поэтому

$$y(c+h)-y(c)=\frac{\partial y(c)}{\partial x_1}h_1+\frac{\partial y(c)}{\partial x_2}h_2+o(h)$$

Этот результат можно написать также в форме

$$y(c+h)-y(c)=\mathcal{D}h+o(h)$$

где равенство $\mathcal{D}h=\frac{\partial y(c)}{\partial x_1}h_1+\frac{\partial y(c)}{\partial x_2}h_2$ определяет $\mathcal{D}$ как непрерывный линейный оператор на пространстве X

Это и доказывает теорему.

б. Методом индукции из а легко получается более общая теорема.

Теорема. Если пространство X есть нормированная прямая сумма подпространств $X_1,\dots,X_n$, функция $y(x)$ $(G\subset X\to Y)$ дифференцируема в некоторой окрестности точки $x=c$ по подпространствам $X_1,\dots,X_n$, причем производные $\frac{\partial y(x)}{\partial x_1},\dots,\frac{\partial y(x)}{\partial x_n}$ непрерывны в точке c, то функция $y(x)$ дифференцируема в точке c по пространству X.

в. Применим теорему б к случаю, когда $X=R_n$, а $X_1,\dots,X_n$ – одномерные подпространства, отвечающие координатным осям. Так как производная по одномерному пространству X_k есть частная производная $\frac{\partial}{\partial x_k}$ (I.46б), теорема б теперь приводит к следующему результату классического анализа:

Теорема. Если функция $y(x)$ $(G\subset R_n\to Y)$ имеет в некоторой окрестности точки $x=c$ частные производные $\frac{\partial y(x)}{\partial x_1},\dots,\frac{\partial y(x)}{\partial x_n}$, непрерывные при $x=c$, то функция $y(x)$ дифференцируема в точке c.

Эта теорема дает широкие достаточные условия дифференцируемости функций $y(x): G\subset R_n\to Y$, поскольку она требует лишь наличия частных производных (по всем переменным), непрерывных в данной точке; такого рода условия иногда проверить легче.

Вместе с тем эту теорему можно сформулировать и как некоторое необходимое и достаточное условие: для существования и непрерывности у функции $y(x): G\subset R_n\to Y$ производной $y'(x)$ в области G необходимы и достаточны существова-

ние и непрерывность в области G частных производных
$\frac{\partial y(x)}{\partial x_1}, \dots, \frac{\partial y(x)}{\partial x_n}$.

г. Производные по одномерным подпространствам. Предыдущая теорема позволяет делать вывод о дифференцируемости функции $y(x)$ на основании ее дифференцируемости по подпространствам $X_1, \dots, X_n$, дающим в прямой сумме все X. Существенно, что их число конечно. Следующее предложение позволяет делать вывод о дифференцируемости функции $y(x)$ на основании ее дифференцируемости по всем одномерным подпространствам. Напомним, что функция $y=y(x)$ дифференцируема в точке $x=c$ по одномерному $R_e=\{te\}, (t\in R_1, e\in X)$, если ее приращение при смещении аргумента вдоль R_e допускает выделение главной линейной части

$$y(c+te)-y(c)=y'_e(c)\cdot te+o(t)$$

где $y'_e(c)$ - линейный оператор на пространстве R_e - производная от $y(x)$ по подпространству R_e.

Теорема. Пусть в области $G\subset X$ заданы векторная функция $y(x)(G\to Y)$ и непрерывная операторная функция $\mathcal{D}(x)(G\to L(X,Y))$. Пусть известно, что функция $y(x)$ в каждой точке $c\in G$ имеет производную $y'_e(c)$ по любому одномерному подпространству R_e, и эта производная действует на любой вектор $h\in R_e$ по формуле

$$y'_e(c)\,h\equiv \mathcal{D}(c)\,h.$$

Тогда функция $y(x)$ дифференцируема в области G и $y'(x)=\mathcal{D}(x)$.

Доказательство. В точке $c \in G$ для заданного $h \in X$ мы имеем

$$y(c+h)-y(c)=y'_e(c)\cdot h+o_e(h)=\mathcal{D}(c)h+o_e(h),$$

и мы должны только показать, что величина $o_e(h)$ является бесконечно малой равномерно по всем h, независимо от их направления. На каждой прямой R_e функция $y(x)$ дифференцируема, и можно применить оценку I.41е:

$$|y(c+h)-y(c)-\mathcal{D}(c)h| \le \sup_{0\le\theta\le 1} |[y'_e(c+\theta h)-y'_e(c)]h| =$$

$$= \sup_{0\le\theta\le 1} |[\mathcal{D}(c+\theta h)-\mathcal{D}(c)]h| \le$$

$$\le \sup_{0\le\theta\le 1} \|\mathcal{D}(c+\theta h)-\mathcal{D}(c)\|\cdot|h| \qquad (I)$$

Теперь для заданного $\varepsilon>0$ найдем $\delta>0$ так, чтобы из $|h|<\delta$ следовало $\|\mathcal{D}(c+h)-\mathcal{D}(c)\|<\varepsilon$; это возможно в силу предположения о непрерывности операторной функции $\mathcal{D}(x)$ в точке c. Тогда, беря любое h, $|h|<\delta$, получаем из (I)

$$|y(c+h)-y(c)-\mathcal{D}(c)h| \le \varepsilon|h|$$

где ε уже не зависит от направления h. Отсюда, как мы указали выше, и следует дифференцируемость функции $y(x)$ при $x=c$, а также и равенство $y'(c)=\mathcal{D}(c)$, что и требовалось.

I.48. Использование теоремы о конечном приращении позволяет иногда устанавливать дифференцируемость сложных функций на основании дифференцируемости более простых.

Пусть M метрическое пространство, V - область в нормированном пространстве Y, и пусть дана функция $\varphi(x,y)$, определенная в $M\times V$, принимающая значения в нормированном пространстве Z, ограниченная и равномерно непрерывная в $M\times V$. Рассмотрим метрические пространства $Y(M)$ и $Z(M)$

(I.16 в) состоящие из всех ограниченных непрерывных функций от $x \in M$, принимающих значения соответственно в Y и в Z; их метризация, указанная в I.16 в, в данном случае может быть задана с помощью норм $\|y(x)\|_{Y(M)} = \sup_{x \in M} |y(x)|_Y$, $\|z(x)\|_{Z(M)} = \sup_{x \in M} |z(x)|_Z$. Совокупность тех $y(x) \in Y(M)$, значения которых лежат в V, мы обозначим естественно через $V(M)$; очевидно, $V(M)$ есть область в пространстве $Y(M)$. Как указано в I.16 д, функция $\Phi(x,y)$ определяет непрерывное отображение $Py(x) = \Phi(x, y(x)) : V(M) \to Z(M)$ по формуле $Fy(x) = \Phi(x, y(x))$. Предположим далее, что у функции $\Phi(x,y)$ имеется производная $\frac{\partial \Phi(x,y)}{\partial y} : Y \to Z$, ограниченная и равномерно непрерывная в $M \times V$. Покажем, что тогда и функция $F(y)$ дифференцируема в области $V(M)$, и найдем выражение ее производной как линейного оператора, действующего из $Y(M)$ в $Z(M)$. При каждом фиксированном x для заданных $y(x) \in V(M)$ и $y(x) + h(x) \in V(M)$ мы имеем

$$\Phi(x, y(x) + h(x)) - \Phi(x, y(x)) = \frac{\partial \Phi(x, y(x))}{\partial y} h(x) + R, \qquad (I)$$

где $\frac{\partial \Phi(x, y(x))}{\partial y} \in L(Y, Z)$ и по I.41е

$$|R| \leq \sup_{0 \leq \Theta(x) \leq 1} \left\| \frac{\partial \Phi(x, y(x) + \Theta(x) h(x))}{\partial y} - \frac{\partial \Phi(x, y(x))}{\partial y} \right\|_{L(Y,Z)} \cdot |h(x)| \leq$$

$$\leq \sup_{\substack{0 \leq \Theta(x) \leq 1 \\ x \in M}} \left\| \frac{\partial \Phi(x, y(x) + \Theta(x) h(x))}{\partial y} - \frac{\partial \Phi(x, y(x))}{\partial y} \right\|_{L(Y,Z)} \cdot \sup_{x \in M} |h(x)|$$

Левая часть в (I) - ограниченная непрерывная функция от x. Первое слагаемое в правой части - также непрерывная и ограниченная функция (по I.16а и I.18 в). Следовательно, и второе слагаемое справа является ограниченной и непрерывной функцией от x. Таким образом, равенство (I) можно трактовать, как равенство в пространстве $Z(M)$. Множитель $\frac{\partial \Phi(x, y(x))}{\partial y}$

при каждом фиксированном $x \in M$ есть (непрерывно зависящий от x) линейный оператор, действующий из Y в Z, поэтому функция $\frac{\partial \phi(x, y(x))}{\partial y}$ может быть трактована, как ограниченный линейный оператор (с нормой, не превосходящей $\sup_{x \in M} \left\| \frac{\partial \phi(x, y(x))}{\partial y} \right\|$), действующий из $Y(M)$ в $Z(M)$.
Поэтому первое слагаемое в правой части (I) линейно по $h \in Y(M)$. Второе слагаемое, как видно из оценки (2) и предположенной равномерной непрерывности $\frac{\partial \phi(x, y)}{\partial y}$ имеет в $Z(M)$ норму порядка $o(\|h\|)$. Отсюда следует, что отображение $F(y)$ дифференцируемо в области $V(M)$ и его дифференциал есть первое слагаемое справа в (I). Можно написать также, что

$$F'(y) = \frac{\partial \phi(x, y(x))}{\partial y} \qquad (3)$$

где правая часть понимается как описанный выше линейный оператор, действующий из $Y(M)$ в $Z(M)$.

Задачи, на которые делались ссылки в тексте, будут приведены (вместе с указаниями и ответами) в отдельном выпуске.

В оформлении этого выпуска принимали участие студенты-практиканты Ю.Раппопорт и В.Творогов. Автор приносит им за это свою живую благодарность.

К О Н Е Ц   П Е Р В О Г О   В Ы П У С К А